Midterm Review Answers

Problem 2

Question:

A researcher is studying the association between a new drug and disease remission across two different hospitals. The data collected is summarized in the following stratified contingency tables:

	Remission (+)	Remission (-)	Total
Drug	40	20	60
Placebo	30	30	60

	Remission (+)	Remission (-)	Total
Drug	50	30	80
Placebo	20	40	60

Solution:

To calculate the **Mantel-Haenszel Common Odds Ratio (MH OR)**, we use the formula:

$$OR_{MH} = \frac{\sum \frac{a_i d_i}{n_i}}{\sum \frac{b_i c_i}{n_i}}$$

where for each stratum i:

- a_i = number of exposed cases (Drug & Remission +)
- b_i = number of exposed non-cases (Drug & Remission -)
- c_i = number of unexposed cases (Placebo & Remission +)
- d_i = number of unexposed non-cases (Placebo & Remission -)
- n_i = total number of observations in the stratum

Step 1: Extract Data from Each Hospital

For **Hospital 1:**

$$a_1 = 40, \quad b_1 = 20, \quad c_1 = 30, \quad d_1 = 30, \quad n_1 = 120$$

For **Hospital 2:**

$$a_2 = 50, \quad b_2 = 30, \quad c_2 = 20, \quad d_2 = 40, \quad n_2 = 140$$

Step 2: Compute the Mantel-Haenszel Components

$$\sum \frac{a_i d_i}{n_i} = \frac{(40 \times 30)}{120} + \frac{(50 \times 40)}{140}$$

$$= \frac{1200}{120} + \frac{2000}{140}$$

$$= 10 + 14.29 = 24.29$$

$$\sum \frac{b_i c_i}{n_i} = \frac{(20 \times 30)}{120} + \frac{(30 \times 20)}{140}$$

$$= \frac{600}{120} + \frac{600}{140}$$

$$= 5 + 4.29 = 9.29$$

Step 3: Compute the Mantel-Haenszel Common Odds Ratio

$$OR_{MH} = \frac{24.29}{9.29} = 2.61$$

Step 4: Interpretation

Since the **Mantel-Haenszel common odds ratio** is **2.61**, this suggests that across both hospitals, patients receiving the drug have **2.61 times higher odds** of experiencing remission compared to those receiving the placebo. This indicates a **strong positive association** between the drug and disease remission.

Problem 3

A researcher is conducting a cohort study to investigate the effect of smoking on lung cancer incidence. The study follows two groups of individuals: Smokers and Non-Smokers. However, since age is a confounding variable, the researcher stratifies the participants into two age groups: Young (≤ 50 years) and Old (> 50 years).

Age Group	Smokers (Cases / Total)	Non-Smokers (Cases / Total)
Young (≤ 50)	50 / 5,000	20 / 6,000
Old (> 50)	200 / 3,000	100 / 4,000

Additionally, the researcher uses a standard population with the following age distribution:

- **Young ($\leq$ 50 years):** 7,000 individuals
- **Old ($>$ 50 years):** 5,000 individuals

Part (A)

Compute the age-stratified risk of lung cancer in both smokers and non-smokers.

Solution: The risk of lung cancer in each stratum is given by:

For Smokers:

- Young ($\leq$ 50 years):
$$\hat{p}_{smoker,young} = \frac{50}{500}$$
- Old ($>$ 50 years):
$$\hat{p}_{smoker,old} = \frac{200}{3000} = 0.0667$$

For Non-Smokers:

- Young ($\leq$ 50 years):
$$\hat{p}_{nonsmoker,young} = \frac{20}{6000} = 0.0033$$
- Old ($>$ 50 years):
$$\hat{p}_{nonsmoker,old} = \frac{100}{4000} = 0.025$$

Part (B)

Perform direct standardization to calculate the adjusted risk for both smokers and non-smokers using the standard population.

Solution: We first define the standard reference distribution:

$$t_{young} = 7000, \quad t_{old} = 5000$$

For Smokers:

$$\hat{p}_{smoker} = \frac{t_{young} \cdot \hat{p}_{young,smoker} + t_{old} \cdot \hat{p}_{old,smoker}}{t_{young} + t_{old}} = 0.0336$$

For Non-Smokers:

$$\hat{p}_{nonsmoker} = \frac{t_{young} \cdot \hat{p}_{young,nonsmoker} + t_{old} \cdot \hat{p}_{old,nonsmoker}}{t_{young} + t_{old}} = 0.0123$$

Part (C)

Compute the adjusted risk difference (RD) and the adjusted risk ratio (RR) between smokers and non-smokers.

Solution:

Risk Difference is defined as

$$\hat{RD} = \hat{p}_{smoker} - \hat{p}_{nonsmoker} = 0.0213.$$

The Risk Ratio is defined as

$$\hat{RR} = \frac{\hat{p}_{smoker}}{\hat{p}_{nonsmoker}} = 2.73.$$

Part (D)

Does smoking increase the risk of lung cancer after adjusting for age?

Solution:

Smokers are 2.73 times more likely to develop lung cancer than non-smokers, even after adjusting for age.

The risk difference of 21.3 per 1,000 suggests that if smoking were eliminated, about 21.3 cases per 1,000 individuals could potentially be prevented in a similar population.