BMI 881 Homework 2

Li Ge

November 6, 2019

This is the link to the homework.

Homework

This is in relation to the paper on predicting nonmelanoma skin cancer, Wang et al. (2019).

1. If a test has sensitivity = 80% and specificity 80% and the prevalence of the disease is 9/100,000, what is the positive predictive value (aka "precision") of the test?

   Answer: 0.036%. (See appendix for details)

2. Suppose sensitivity = specificity. What would they have to be to achieve positive predictive value = 50% when prevalence is 9/100,000?

   Answer: 99.991% (See appendix for details)

3. Comment on these results in relation to the precision values provided in Table 2 of Wang et al. (2019).

   In the study, a total of 1829 patients with nonmelanoma skin cancer (NMSC) as their first diagnosed cancer and 7665 random controls without cancer were included in the analysis. Their best reported $Sn = 83.1\%, Sp = 82.3\%, Precision = 57.1\%$.

   However, the NMSC incidence among Asian individuals is 2.3 to 9.2 per 100,000 population. If we apply their methods to the real data, the precision will be strikingly low like the case in question 1. The conclusion is that their study design didn't reflect the real prevalence of the disease. Thus, the reported precision is not at all reliable. The proposed method is questionable at the best since the results are based on fictional dataset.

Appendix

Table 1: Confusion Matrix

Prediction \ Truth	Positive	Negative	Sum
Positive	$np \times Sn$	$n(1-p) \times (1-Sp)$	$np \times Sn + n(1-p) \times (1-Sp)$
Negative	$np \times (1-Sn)$	$n(1-p) \times Sp$	$np \times (1-Sn) + n(1-p) \times Sp$
Sum	np	$n(1-p)$	np

- n: # of samples

- p: prevalence rate

- Sn: sensitivity

- Sp: specificity

$$Precision = \frac{TP}{TP + FP} = \frac{np \times Sn}{np \times Sn + n(1-p) \times (1-Sp)}$$
$$= \frac{p \times Sn}{p \times Sn + (1-p) \times (1-Sp)}$$

1. Solve question 1 by plugging in $p = 0.00009, Sn = Sp = 0.8$.

2. Solve question 2 by solving

$$Precision = 0.5$$
$$p = 0.0009$$
$$Sn = Sp$$