M3S12 BIOSTATISTICS - EXERCISE 2

MEASURES OF EFFECT

1. Show, by considering the probability of exposure, that the *Risk Difference* in an epidemiological study, $\pi_1 - \pi_0$, can be an important measure of effect for health care provision.

2. In a closed cohort study, derive the formulae for maximum likelihood estimates for the following parameters

   (i) $\pi_1 = P(F|E)$

   (ii) $\pi_0 = P(F|E')$

   (iii) $\theta = P(E)$

   (iv) $\phi = P(F)$

   where as usual E and E' correspond to exposure/non-exposure and F and F' correspond to presence/absence of disease in the cohort.

3. For each of the following 2×2 tables, compute the risk difference, relative risk, and odds ratio assuming that the results are derived from closed cohort studies.

(i)

	E	E'
F	210	120
F'	90	180

(ii)

	E	E'
F	230	120
F'	70	80

(iii)

	E	E'
F	170	80
F'	130	220

(iv)

	E	E'
F	120	130
F'	180	170

4. Let X be a binary potential risk factor, and consider the following pairs of 2×2 tables. Decide whether X is a risk factor for the disease, and whether X is related to exposure.

(i) For $X = 0$:

	E	E'
F	70	80
F'	30	120

For $X = 1$:

	E	E'
F	160	40
F'	40	60

(ii) For $X = 0$:

	E	E'
F	90	60
F'	10	40

For $X = 1$:

	E	E'
F	80	20
F'	120	180

5. The data below arise from a case-control study carried out to investigate the relationship between oesophagel cancer (diseased or not, F, F') and alcohol consumption ($\geq 80g/ < 79g$ on average per day for exposed/unexposed groups, E, E'). It is known that the proportion of exposed individuals in the general population is $\theta = 0.40$.

	E	E'
F	96	104
F'	109	666

Recall that the *odds ratio* is defined by

$$\psi = \frac{\pi_1/\pi_0}{(1-\pi_1)\,/\,(1-\pi_0)} = \frac{P(F|E)/P(F|E')}{P(F'|E)/P(F'|E')} = \frac{P(E|F)/P(E'|F)}{P(E|F')/P(E'|F')}.$$

Find the maximum likelihood estimates and associated standard errors for

(i) π_1 and π_0

(ii) Risk difference $RD : \delta = \pi_1 - \pi_0$. Interpret your answer.

(iii) Can you find a confidence interval for ψ ?

Recall that the *standard error* of an estimator T of parameter θ is

$$s.e.\left(T;\theta\right) = \sqrt{Var_{f_{T|\theta}}\left[T\right]} = s_e\left(\theta\right)$$

for some function s_e, and the *estimated standard error* is

$$e.s.e\left(T\right) = s_e\left(\widehat{\theta}\right)$$

where $\widehat{\theta}$ is the computed maximum likelihood estimate (that is, the observed value of T).