生物統計 HW3 解答

#1: 判斷 independence

Approximately 50% of the population is male, 68% drinks to some extent, and 38.5% drinks and is male. Given that a randomly selected individual is male, find the probability that he drinks. Is a person's drinking status independent of gender?

Sol):

Let $A \equiv male, B \equiv drinks$

$\Pr(A) = 0.5, \Pr(\mathrm{B}) = 0.68, \Pr(A \cap B) = 0.385$

(a) 求 $\Pr(\mathrm{B} \mid A)$ =?

$$\Pr(\mathrm{B} \mid A) = \frac{\Pr(\mathrm{A} \cap \mathrm{B})}{\Pr(\mathrm{A})} = \frac{0.385}{0.5} = 0.77$$

(b) A 和 B 是否獨立?

$\because \Pr(\mathrm{B} \mid A) = 0.77 \neq 0.68 = \Pr(\mathrm{B})$

$\therefore$ A 和 B 不獨立

#2: Case-control data 實例

In a study of the relationship between the regular use of hair dye and the development of leukemia, 577 leukemia patients and 1245 persons free from the disease (controls) were selected and questioned concerning their use of hair dye. Forty-three patients and 55 controls claimed to have had significant exposure to hair dye. (Based on information found in Kenneth Cantor et al., "Hair Dye Use and Risk of Leukemia and Lymphoma," *American Journal of Public Health,* vol. 78, no. 5, May 1988, pp. 570–571.)

TABLE 3.12

		Leukemia present	
		Yes	No
Use hair dye	Yes	43	55
	No		
		577 (Fixed)	1245 (Fixed)

(a)

Complete Table 3.12. (4分)

Sol):

		Leukemia present	
		Yes	No
Use hair dye	Yes	43	55
	No	534	1190
		577(Fixed)	1245(Fixed)

(b)

In this case, is it possible to approximate relative risk using the definition given in this section? Explain. (5分)

Sol):

No；因資料來源不同，不同 column 不可相加。

(c)

Some idea of the impact of hair dye use can be obtained by considering the ratio (11分)

$$\frac{P[E \mid D]}{P[E \mid D']}$$

where E is the event that the individual was exposed to risk and D the event that leukemia is present. Can each of the conditional probabilities involved in this ratio be estimated? If so, evaluate and interpret the ratio.

Sol):

Yes；有白血病的人使用染髮劑的 odds 是沒有白血病的人的 1.68 倍。

$$\frac{\frac{43}{577}}{\frac{55}{1245}} = \frac{1}{\frac{577*55}{1245*43}} = \frac{1}{0.5927} = 1.68$$

#3:

In 1985, many Asian children were adopted by American families. Some of these children had been exposed to the hepatitis B virus and were capable of transmitting the virus to others. In a study of the risk involved, the data of Table 3.11 were obtained. Cell entries represent the number of close family members found with the virus, and all row and column totals are random. Approximate the relative risk. (Based on information found in Andrew Friede et al., "Transmission of Hepatitis B Virus from Adopted Asian Children to Their American Families," *American Journal of Public Health,* vol. 78, no. 1, January 1988, pp. 26–29.)

TABLE 3.11

		Virus present	
		Yes	**No**
Exposed to risk	Yes	7	70
	No	4	228

(a)

求 因領養而感染B型肝炎病毒之 relative risk and odds ratio (16分)

Sol):

$$RR = \frac{P(\mathrm{D}|\mathrm{E})}{P(D|\mathrm{E}')} = \frac{\frac{7}{77}}{\frac{4}{232}} = 5.27$$

$$OR = \frac{7\times 228}{4\times 70} = 5.7$$

(b)

兩者有相關性嗎？(指感染 virus 是否來自領養孩童) (4分)

Sol):

因為 $RR = 5.27 > 1$，所以代表感染與領養孩童有相關。

#4: Bayes Theorem

Statistics indicate that the probability that a mother will die during childbirth in the United States is .00022. If the mother is not black, the probability of death is .00017, whereas the figure is .00064 if she is black. Assume that 10% of the births recorded are to blacks.

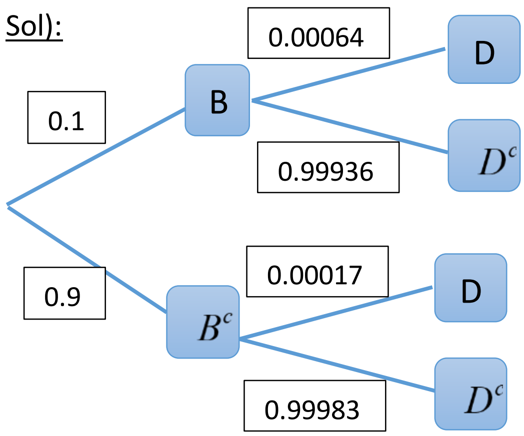
Let $D \equiv die, B \equiv black$

$$\Pr(\mathrm{D}) = 0.00022, \Pr(\mathrm{D} \mid \mathrm{B}^c) = 0.00017, \Pr(\mathrm{D} \mid \mathrm{B}) = 0.00064, \Pr(B) = 0.1$$

(a)

Draw a tree indicating these probabilities, and find the path probabilities for each of the four paths. (Let *D* denote the event that the mother dies and *B* the event that the mother is black.)

Sol):



(b)

Use the tree of part *a* to find the probability that a mother who dies in childbirth is black.

Sol):

$$\Pr(\mathrm{B} \mid \mathrm{D}) = \frac{\Pr(\mathrm{B} \cap \mathrm{D})}{\Pr(\mathrm{D})} = \frac{0.1 \times 0.00064}{0.00022} = 0.2909$$

(c)

Using Bayes' theorem, find the probability that a mother who dies in childbirth is black, and compare your answer to that obtained in part *b*.

Sol):

題目有規定使用貝式定理，故分母需使用先驗機率做計算。

$$\Pr(\mathrm{B} \mid \mathrm{D}) = \frac{\Pr(\mathrm{B} \cap \mathrm{D})}{\Pr(\mathrm{D} \mid \mathrm{B})\Pr(\mathrm{B}) + \Pr(\mathrm{D} \mid \mathrm{B}^c)\Pr(\mathrm{B}^c)} = \frac{0.1 \times 0.00064}{0.1 \times 0.00064 + 0.9 \times 0.00017} = 0.2949$$

(b)&(c)答案近似，只在分母小數進位的差別。

<u>#5:</u> Discrete random variable

The following table shows the density for the random variable X, the number of persons seeking emergency room treatment unnecessarily per day in a small hospital.

x	0	1	2	3	4	5
$f(x)$	.01	.1	.3	.4	.1	?

(a)

Find $f(5)$. What probability does this represent in the context of this problem?

<u>Sol):</u>

$$\sum_{x=0}^{5} f(x) = 1$$

$$\rightarrow 0.01 + 0.1 + 0.3 + 0.4 + 0.1 + f(5) = 1$$

$$\rightarrow f(5) = 0.09$$

The probability of 5 people seeking emergency room treatment unnecessarily per day is 0.09.

(b)

Find $P[X \leq 2]$. Interpret this probability in the context of this problem.

<u>Sol):</u>

$$\Pr(X \leq 2) = \sum_{x=0}^{2} f(x) = 0.01 + 0.1 + 0.3 = 0.41$$

The probability of people less than or equal to 2 seeking emergency room treatment unnecessarily per day is 0.41.

(c)

Find $P[X < 2]$.

<u>Sol):</u>

$$\Pr(X < 2) = \sum_{x=0}^{1} f(x) = 0.01 + 0.1 = 0.11$$

(d)

Find $P[X > 3]$

<u>Sol):</u>

$$\Pr(X > 3) = \sum_{x=4}^{5} f(x) = 0.1 + 0.09 = 0.19$$

(e)

Compute $E(X)$

Sol):

$$E(\mathrm{X})=\sum_{x=0}^{5} x f(x)=1\times 0.1+2\times 0.3+3\times 0.4+4\times 0.1+5\times 0.09=2.75$$

(f)

compute $Var(X)=\sigma^2$ and σ

Sol):

$$E(\mathrm{X}^2)=\sum_{x=0}^{5} x^2 f(x)=1\times 0.1+2^2\times 0.3+3^2\times 0.4+4^2\times 0.1+5^2\times 0.09=8.75$$

$$\rightarrow \sigma^2=Var(X)=E(\mathrm{X}^2)-[\mathrm{E(X)}]^2=8.75-2.75^2=1.1875$$

$$\rightarrow \sigma=1.0897$$