STT 465 HW1

[redacted]

Question 1. In population I $\theta = 0.9$, compute and report the prevalence of disease in this population.

Solution.

$$
\begin{aligned}
P(D=1) &= P(D=1|G=AA)P(G=AA) + P(D=1|G=AB)P(AB) + P(D=1|G=BB)P(BB) \\
&= (0.0)\theta^2 + (0.1)(2)\theta(1-\theta) + 0.9(1-\theta)^2 \\
&= (0.1)(2)(0.9)(1-0.9) + 0.9(1-0.9)^2 \\
P(D=1) &= 0.027
\end{aligned}
$$

Question 2. An individual in population I has developed the disease, what is the probability that the genotype of that individual is AA, AB, or BB?

Solution.

1. AA

$$P(G_i = AA|D=1) = \frac{P(D=1 \ \& \ G_i = AA)}{P(D=1)} = \frac{P(D=1|G_i=AA)P(G_i=AA)}{P(D=1)} = \frac{0(0.9)}{0.027} = 0$$

2. AB

$$P(G_i = AB|D=1) = \frac{P(D=1 \ \& \ G_i = AB)}{P(D=1)} = \frac{P(D=1|G_i=AB)P(G_i=AB)}{P(D=1)} = \frac{0.1(2*0.9*(1-0.9))}{0.027} = \frac{2}{3}$$

3. BB

$$P(G_i = BB|D=1) = \frac{P(D=1 \ \& \ G_i = BB)}{P(D=1)} = \frac{P(D=1|G_i=BB)P(G_i=BB)}{P(D=1)} = \frac{0.9(1-0.9)^2}{0.027} = \frac{1}{3}$$

Question 3. In population II the frequency of allele A is 0.95. An individual is healthy, what is the probability that this individual comes from population I (assume that, a priori, individuals are equally likely to come from population 1 or 2).

Solution. Let N be the population that the individual comes from. P(N = II)=P(N = I)=0.5. From question 1,

$$P(D=1|N=I) = 0.027$$

For population II,

$$
\begin{aligned}
P(D=1|N=II) &= P(D=1|G=AA)P(G=AA) + P(D=1|G=AB)P(AB) + P(D=1|G=BB)P(BB) \\
&= (0.0)\theta^2 + (0.1)(2)\theta(1-\theta) + 0.9(1-\theta)^2 \\
&= (0.1)(2)(0.95)(1-0.95) + 0.9(1-0.95)^2 \\
P(D=1|N=II) &= 0.01175
\end{aligned}
$$

Therefore,

$$P(N = I|D = 0) = \frac{P(N = I \;\&\; D = 0)}{P(D = 0)} = \frac{P(D = 0|N = I)P(N = I)}{P(D = 0|N = I)P(N = I) + P(D = 0|N = II)P(N = II)}$$

$$P(N = I|D = 0) = \frac{(1 - P(D = 1|N = I))P(N = I)}{(1 - P(D = 1|N = I))P(N = I) + (1 - P(D = 1|N = II))P(N = II)}$$

$$P(N = I|D = 0) = \frac{(1 - 0.027) * 0.5}{(1 - 0.027) * 0.5 + (1 - 0.01175) * 0.5} = 0.496112$$

Question 4. Provide frequencies for the joint distribution of the two Bernoulli random variables (i.e., probabilities in a 2x2 contingency table) that:

1. satisfy IID (identically and independently distributed),
2. satisfy exchangibility but not IID
3. do no satisfy IID and are not exchangable

Explain your reasoning.

Solution.

1. The below table satisfies IID.

X

Y		0	1	P(Y=y)
	0	0.25	0.25	0.5
	1	0.25	0.25	0.5
	P(X=x)	0.5	0.5	

This satisfies IID because (1) both X and Y have the same marginal distribution and (2) their joint distribution is the product of their marginal distributions.

$$\begin{aligned}
(1)\;& P(X = 1) = P(Y = 1) = 0.5 \\
& P(X = 0) = P(Y = 0) = 0.5 \\
(2)\;& P(X = 1, Y = 1) = P(Y = 1)P(X = 1) = 0.5(0.5) = 0.25 \\
& P(X = 0, Y = 1) = P(Y = 1)P(X = 0) = 0.5(0.5) = 0.25 \\
& P(X = 0, Y = 0) = P(Y = 0)P(X = 0) = 0.5(0.5) = 0.25 \\
& P(X = 1, Y = 0) = P(Y = 0)P(X = 1) = 0.5(0.5) = 0.25
\end{aligned}$$

2. The below table satisfies exchangeability but not IID.

X

Y		0	1	P(Y=y)
	0	0.1	0.3	0.4
	1	0.3	0.3	0.6
	P(X=x)	0.4	0.6	

This does not satisfy IID because:

$$P(X=0,Y=0) \neq P(X=0)P(Y=0)$$
$$0.1 \neq 0.4(0.4) = 0.16$$

This does satisfy exchangeability becasue:

$$(1)\ P(X=0,Y=0) = P(X=0,Y=0) = 0.1$$
$$(2)\ P(X=1,Y=0) = P(X=0,Y=1) = 0.3$$
$$(3)\ P(X=0,Y=1) = P(X=1,Y=0) = 0.3$$
$$(4)\ P(X=1,Y=1) = P(X=1,Y=1) = 0.3$$

3. The below table does not satisfy either exchangeability or IID.

		X = 0	X = 1	P(Y=y)
Y	0	0.05	0.25	0.3
	1	0.35	0.35	0.7
	P(X=x)	0.4	0.6	

This does not satisfy IID because:

$$P(X=0,Y=0) \neq P(X=0)P(Y=0)$$
$$0.1 \neq 0.4(0.3) = 0.12$$

This does not satisfy exchangeability because:

$$P(X=1,Y=0) \neq P(X=0,Y=1)$$
$$0.25 \neq 0.35$$

Question 5. Consider a system of three Bernoulli random variables (X,Y,Z). P(Z=1)=0.6. The following tables give the conditional distributions $P(X,Y|Z)$

$P(X,Y|Z=0)$

	Y=0	Y=1
X=0	0.06	0.24
X=1	0.14	0.56

$P(X,Y|Z=1)$

	Y=0	Y=1
X=0	0.12	0.28
X=1	0.18	0.42

5.a) Are (X,Y) conditionally independent?
5.b) Are (X,Y) independent?
5.c) Are (X,Y) exchangeable?

Solution. First I compute the conditional distributions of X and Y on Z

$P(X, Y|Z = 0)$ | $P(X, Y|Z = 1)$

	Y=0	Y=1	$P(X = x\|Z = 0)$		Y=0	Y=1	$P(X = x\|Z = 1)$
X=0	0.06	0.24	0.3	X=0	0.12	0.28	0.4
X=1	0.14	0.56	0.7	X=1	0.18	0.42	0.6
$P(Y = y\|Z = 0)$	0.2	0.8		$P(Y = y\|Z = 1)$	0.3	0.7	

5.a) (X,Y) are conditionally independent if $P(X = x, Y = y|Z = x) = P(X = x|Z = z)P(Y = y|Z = z)$ for all $x, y, z \in \{0, 1\}$.
For $Z = 0$,

$$P(X = 0, Y = 0) = 0.06 = (0.3)(0.2) = P(X = 0|Z = 0)P(Y = 0|Z = 0) \checkmark$$
$$P(X = 0, Y = 1) = 0.24 = (0.3)(0.8) = P(X = 0|Z = 0)P(Y = 1|Z = 0) \checkmark$$
$$P(X = 1, Y = 0) = 0.14 = (0.7)(0.2) = P(X = 1|Z = 0)P(Y = 0|Z = 0) \checkmark$$
$$P(X = 1, Y = 1) = 0.56 = (0.7)(0.8) = P(X = 1|Z = 0)P(Y = 1|Z = 0) \checkmark$$

For Z=1,

$$P(X = 0, Y = 0) = 0.12 = (0.4)(0.3) = P(X = 0|Z = 1)P(Y = 0|Z = 1) \checkmark$$
$$P(X = 0, Y = 1) = 0.28 = (0.4)(0.7) = P(X = 0|Z = 1)P(Y = 1|Z = 1) \checkmark$$
$$P(X = 1, Y = 0) = 0.18 = (0.6)(0.3) = P(X = 1|Z = 1)P(Y = 0|Z = 1) \checkmark$$
$$P(X = 1, Y = 1) = 0.42 = (0.6)(0.7) = P(X = 1|Z = 1)P(Y = 1|Z = 1) \checkmark$$

Therefore, (X,Y) are conditionally independent because $P(X = x, Y = y|Z = x) = P(X = x|Z = z)P(Y = y|Z = z)$ for all $x, y, z \in \{0, 1\}$.

5.b) To find if (X,Y) are independent, we must first compute the joint distribution of X,Y:

Expressions in the left-hand-side are the joint prob. of X and Y. Those in the middle and in the right-hand-side are the joint probabilities of X and Y given Z. The two are not equal. The middle and right-hand side are correct.

$$\begin{aligned}
P(X = 0, Y = 0) &= P(X = 0, Y = 0|Z = 0)P(Z = 0) + P(X = 0, Y = 0|Z = 1)P(Z = 1) \\
&= (0.06)(1 - 0.6) + (0.12)(0.6) = 0.096 \\
P(X = 0, Y = 1) &= P(X = 0, Y = 1|Z = 0)P(Z = 0) + P(X = 0, Y = 1|Z = 1)P(Z = 1) \\
&= (0.24)(1 - 0.6) + (0.28)(0.6) = 0.264 \\
P(X = 1, Y = 0) &= P(X = 1, Y = 0|Z = 0)P(Z = 0) + P(X = 1, Y = 0|Z = 1)P(Z = 1) \\
&= (0.14)(1 - 0.6) + (0.18)(0.6) = 0.164 \\
P(X = 1, Y = 1) &= P(X = 1, Y = 1|Z = 0)P(Z = 0) + P(X = 1, Y = 1|Z = 1)P(Z = 1) \\
&= (0.56)(1 - 0.6) + (0.42)(0.6) = 0.476
\end{aligned}$$

	Y=0	Y=1	P(X=x)
X=0	0.096	0.264	0.36
X=1	0.164	0.476	0.64
P(Y=y)	0.26	0.74	

For X,Y to be independent P(X,Y)=P(X)P(Y) for all X,Y.

$$P(X = 0, Y = 0) = 0.096$$

$$P(X = 0)P(Y = 0) = (0.36)(0.26) = 0.936 \neq 0.096 = P(X = 0, Y = 0)$$

Since $P(X = 0, Y = 0) \neq P(X = 0)P(Y = 0)$, (X,Y) are not independent.

5.c) For X,Y to be exchangeable $P(X = i, Y = j) = P(X = j, Y = i)$ for all $i, j \in \{0, 1\}$.

$$P(X = 0, Y = 1) = 0.264$$
$$P(X = 1, Y = 0) = 0.164 \neq 0.264$$

Since $P(X = 0, Y = 1) \neq P(X = 1, Y = 0)$, X and Y are not exchangeable.