Problem Set 2

Q1.

a) 9.8%

The probability that a randomly selected engineer programs in both Java and C++ is:

$P(J \cap C{+}{+}) = P(J) \times P(C{+}{+} \mid J) = 0.35 \times 0.28 = 0.098$

b)24.5%

The conditional probability that a randomly selected engineer programs in Java given that they program in C++ is:

$P(J \mid C{+}{+}) = P(J \cap C{+}{+}) / P(C{+}{+}) = 0.098 / 0.40 = 0.245$

Q2.

a) 99.76%

A website gives a visitor five independent CAPTCHA tests; failing any one flags the visitor as a robot. Humans pass each test with probability 0.95, robots with probability 0.3, and robots form 5% of visitors.

If the visitor is actually a robot, the probability they are flagged is:

$1 - (0.3)^5 = 1 - 0.00243 = 0.99757 \approx$ **99.76%**

b) 22.62%

If the visitor is human, the probability they are flagged is:

$1 - (0.95)^5 = 1 - 0.77378 = 0.22622 \approx$ **22.62%**

c) 18.78%

Using Bayes' theorem, the overall probability of being flagged is:

$0.05(0.99757) + 0.95(0.22622) \approx 0.26579$

Thus, the probability a flagged visitor is actually a robot is:

$(0.05 \times 0.99757) / 0.26579 \approx 0.1878 =$ **18.78%**

d) 2.14%

If a visitor is human, the probability they pass exactly three of five tests is:

$C(5,3) \times (0.95^3) \times (0.05^2) = 10 \times 0.857375 \times 0.0025 \approx 0.02143 =$ **2.14%**

e) 2.70%

For a visitor of unknown identity, the probability of exactly three passes is:

$0.05[C(5,3)(0.3^3)(0.7^2)] + 0.95[C(5,3)(0.95^3)(0.05^2)]$

$= 0.05(0.1323) + 0.95(0.02143) \approx 0.02698 =$ **2.70%**

Q3. 82.35%

All computers run either OS W or X. A computer running W is twice as likely to get infected as one running X. If 70% of all computers run W and 30% run X, let the infection rate for X be r and for W be 2r.

The fraction of infected computers that run W is:

$(0.7 \times 2r) / (0.7 \times 2r + 0.3 \times r) = 1.4 / 1.7 = 14 / 17 \approx 0.8235$

Thus, about **82.35%** of infected computers run W.

Q4. 58.14%

The referee chooses one of three coins (with head probabilities 0.1, 0.5, 0.9) at random and tosses it three times. The sequence H, T, H has likelihood $p^2(1-p)$.

For $p=0.1 \rightarrow 0.009$, for $p=0.5 \rightarrow 0.125$, for $p=0.9 \rightarrow 0.081$.

Using Bayes' theorem:

$P(\text{fair coin} \mid H,T,H) = 0.125 / (0.009 + 0.125 + 0.081) = 0.125 / 0.215 \approx 0.5814$

Thus, the probability is about 58.14%.

Q5. $q = P(\text{non-spam} \mid \text{GOOD}) = p / (0.9p + 0.1)$

$q = p / (0.9p + 0.1)$

Let p be the prior probability an email is non-spam. The filter marks an email GOOD if it is non-spam (always) or spam with a 10% bug.

$P(\text{GOOD}) = p + 0.1(1 - p) = 0.9p + 0.1$

$P(\text{non-spam} \cap \text{GOOD}) = p$

Thus, **$q = P(\text{non-spam} \mid \text{GOOD}) = p / (0.9p + 0.1)$**

Compare q and p:

$q - p = [p / (0.9p + 0.1)] - p = (0.9p(1 - p)) / (0.9p + 0.1) \geq 0$ for $0 \leq p \leq 1$

So $q > p$, meaning a GOOD email is more likely to be non-spam than average.

Q6.

a) 5.88%

Given the Ace of Spades is one of the two drawn cards, the probability both cards are Aces is $3/51 = 1/17 \approx 5.88\%$.

b) Not independent

The unconditional probability that both cards are Aces is $C(4,2) / C(52,2) = 6/1326 = 1/221 \approx 0.452\%$.

Since $P(E|F) = 1/17 \neq P(E)$, the events are not independent.

c) 3.03%

The probability of at least one Ace is $1 - C(48,2)/C(52,2) = 198/1326 = 33/221$.

Thus, $P(E|G) = (1/221) / (33/221) = 1/33 \approx 3.03\%$.

Q7.

a) Yes

Given $P(G) = 0.6$, $P(T1 \mid G) = 0.7$, $P(T2 \mid G) = 0.9$, and $P(T1 \text{ and } T2 \mid G) = 0.63$, we check conditional independence by comparing $P(T1 \mid G) \times P(T2 \mid G) = 0.7 \times 0.9 = 0.63$ to $P(T1 \text{ and } T2 \mid G) = 0.63$. They are equal, so T1 and T2 are conditionally independent given G.

b) Yes

If a subject does not have G, they express neither T1 nor T2, so $P(T1 \mid G^c) = 0$ and $P(T2 \mid G^c) = 0$ and $P(T1 \text{ and } T2 \mid G^c) = 0$. The product $P(T1 \mid G^c) \times P(T2 \mid G^c) = 0$ equals $P(T1 \text{ and } T2 \mid G^c) = 0$, so T1 and T2 are (trivially) conditionally independent given G^c.

c) 0.42

$P(T1) = P(G)P(T1 \mid G) + P(G^c)P(T1 \mid G^c) = 0.6 \times 0.7 + 0.4 \times 0 = 0.42$.

d) 0.54

$P(T2) = P(G)P(T2 \mid G) + P(G^c)P(T2 \mid G^c) = 0.6 \times 0.9 + 0.4 \times 0 = 0.54$.

e) No

$P(T1 \text{ and } T2) = P(G) \times P(T1 \text{ and } T2 \mid G) = 0.6 \times 0.63 = 0.378$, while $P(T1)P(T2) = 0.42 \times 0.54 = 0.2268$. Since these differ, T1 and T2 are not independent marginally.

Q8.

a) 2/3

Claire has blue eyes, so both parents must carry a blue gene (they are each genotype B b). For parents B b × B b, children's genotypes are: BB (1/4), B b (1/2), b b (1/4). William has brown eyes (not b b), so conditional on brown his genotype is BB or B b with relative probabilities 1/4 : 1/2. Thus P(William is B b | brown) = (1/2) / (3/4) = 2/3. So the probability William possesses a blue-eyed gene is 2/3.

b) 1/3

William's wife is blue-eyed (b b), so she always contributes a blue gene. William will contribute a blue gene only if he is a carrier B b (probability 2/3) and then transmits the blue gene with probability 1/2. Hence the child's probability of blue eyes = (2/3) × (1/2) = 1/3.

Q9. Not done