MSE Hand in 5 - Solution

You can solve the problems any way you like, but the answers must be submitted individually in WiseFlow. Submission by the end of class. You must upload your calculations in WiseFlow in the documentation flow. This Flow is opened 30 minutes after the assignment Flow closes. You must upload one pdf with all your calculations to each assignment. The assignments must be stated in numerical order (Assignment 1, Assignment 2, etc.).

Assignment 1

The following is called the **Newton-Pepys problem**. Isaac Newton was consulted about the following problem by Samuel Pepys, who wanted the information for gambling purposes. Find the probability (four decimal precision) of each of the three events and determine which has the highest probability? State the probabilities with four decimal precision.

- A: At least one 6 appears when 6 fair dice are rolled.

$$\begin{aligned} P(A) &= 1 - P(\text{no 6's appear}) \\ &= 1 - \left(\frac{5}{6}\right)^6 \\ &\approx \underline{\underline{0.6651}} \end{aligned}$$

- B: At least two 6's appear when 12 fair dice are rolled.

$$\begin{aligned} P(B) &= 1 - P(\text{no 6's appear}) - P(\text{exactly one 6 appears}) \\ &= 1 - \left(\frac{5}{6}\right)^{12} - \binom{12}{1}\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^{11} \\ &\approx \underline{\underline{0.6187}} \end{aligned}$$

- C: At least three 6's appear when 18 fair dice are rolled.

$$\begin{aligned} P(C) &= 1 - P(\text{no 6's appear}) - P(\text{exactly one 6 appears}) - P(\text{exactly two 6's appear}) \\ &= 1 - \left(\frac{5}{6}\right)^{18} - \binom{18}{1}\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^{17} - \binom{18}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{16} \\ &\approx \underline{\underline{0.5973}} \end{aligned}$$

Assignment 2

Computer keyboard failures are due to faulty electrical connects (12%) or mechanical defects (88%). Mechanical defects are related to loose keys (27%) or improper assembly (73%). Electrical connect defects are caused by defective wires (35%), improper connections (13%), or poorly welded wires (52%). State the probabilities with four decimal precision.

(a) Find the probability that a failure is due to loose keys.

$$\begin{aligned} P(\text{loose keys}) &= P(\text{mechanical defect}) \cdot P(\text{loose keys} \mid \text{mechanical defect}) \\ &= 0.88 \cdot 0.27 \\ &\approx \underline{\underline{0.2376}} \end{aligned}$$

(b) Find the probability that a failure is due to improperly connected or poorly welded wires. Define:
ECD = electrical connect defect, IC = improperly connected, PW = poorly welded wires
Then the probability is:

$$\begin{aligned} P(\text{IC or PW}) &= P(\text{ECD}) \cdot P(\text{IC or PW} \mid \text{ECD}) \\ &= 0.12 \cdot (0.13 + 0.52) \\ &\approx \underline{\underline{0.0780}} \end{aligned}$$

Assignment 3

Assume that $P(B \mid A) = 0.3$, $P(A) = 0.9$, $P(B \mid A^c) = 0.5$.

(a) Find the total probability for B, $P(B)$. State your answer with two decimal precision.
We can use the law of total probability to find $P(B)$:

$$\begin{aligned} P(B) &= P(B \mid A) \cdot P(A) + P(B \mid A^c) \cdot P(A^c) \\ &= 0.3 \cdot 0.9 + 0.5 \cdot 0.1 \\ &\approx \underline{\underline{0.32}} \end{aligned}$$

(b) Find the probability for A given B, $P(A \mid B)$. State your answer with four decimal precision. We can use Bayes' theorem to find $P(A \mid B)$:

$$\begin{aligned} P(A \mid B) &= \frac{P(B \mid A) \cdot P(A)}{P(B)} \\ &= \frac{0.3 \cdot 0.9}{0.32} \\ &\approx \underline{\underline{0.8438}} \end{aligned}$$

Assignment 4

Even more planes! The probability that a regularly scheduled flight departs on time is 0.81; the probability that it arrives on time is 0.80; and the probability that it departs and arrives on time is 0.76. Find the probability that a plane arrives on time, given that it did not depart on time. State your answer with two decimal precision.
Define: D = departs on time, A = arrives on time
Then the probability is:

$$\begin{aligned} P(A \mid D^c) &= \frac{P(A \cap D^c)}{P(D^c)} \\ &= \frac{P(A) - P(A \cap D)}{1 - P(D)} \\ &= \frac{0.80 - 0.76}{1 - 0.81} \\ &\approx \underline{\underline{0.21}} \end{aligned}$$

Assignment 5

An email user would like to predict what the probability of a given mail being spam is, given that the email starts with the words "Dear Friend". He knows the following:

- Of all the emails he receives, 10% are spam. So $P(\text{spam}) = 0.1$.

- Of all the spam emails he receives, 15% starts with the words "Dear Friend". From this he concludes that $P(\text{"Dear Friend"} \mid \text{spam}) = 0.15$.
- Of all the non-spam emails he receives, 1% starts with the words "Dear Friend". From this he concludes that $P(\text{"Dear Friend"} \mid \text{not spam}) = 0.01$.

Find the following probabilites and state your answers with three decimal precision.

(a) Calculate the total probability that an email he receives starts with "Dear Friend", $P(\text{ "Dear Friend"})$.

$$\begin{aligned} P(\text{"Dear Friend"}) &= P(\text{"Dear Friend"} \mid \text{spam}) \cdot P(\text{spam}) + P(\text{"Dear Friend"} \mid \text{not spam}) \cdot P(\text{not spam}) \\ &= 0.15 \cdot 0.1 + 0.01 \cdot 0.9 \\ &\approx \underline{\underline{0.024}} \end{aligned}$$

(b) Calculate the probability that an email is spam given that it starts with the words "Dear Friend", P (spam—"Dear Friend").

$$\begin{aligned} P(\text{spam} \mid \text{"Dear Friend"}) &= \frac{P(\text{"Dear Friend"} \mid \text{spam}) \cdot P(\text{spam})}{P(\text{"Dear Friend"})} \\ &= \frac{0.15 \cdot 0.1}{0.024} \\ &\approx \underline{\underline{0.625}} \end{aligned}$$

Assignment 6

A recent study made by the department of education here in Denmark asked students in three different areas what was the most important thing in school: making good grades ('Grades), having a high learning outcome ('Learning'), or having a good social life ('Social'). Students from rural, suburban, and urban areas were surveyed. A total of 478 students participated in the survey and 149 were from rural areas and of these 50 answered 'Learning'. A total of 247 students answered 'Grades' and of these 103 lived in urban areas. 42 of the 141 who answered 'Learning' lived in a suburban area. 178 of the respondents lived in an urban area. There were 30 more respondents living in a suburban area who answered 'Grades' than there were respondents living in rural areas who answered 'Grades'. State all probabilities with two decimal precision.

(a) Based on the above information, create a 3×3 contingency table. Please place area of living on the horizontal axis.

	Rural	Suburban	Urban	Total
Grades	57	87	103	247
Learning	50	42	149	141
Social	42	22	26	90
Total	149	151	178	478

(b) What is the probability that a randomly chosen student was from a suburban area and thought having a good social life was most important?

$$\begin{aligned} P(\text{suburban} \cap \text{social}) &= \frac{22}{478} \\ &\approx \underline{\underline{0.05}} \end{aligned}$$

(c) What is the probability that a randomly chosen student was from a suburban area, given the student thought having a good social life was most important?

$$\begin{aligned} P(\text{suburban} \mid \text{social}) &= \frac{22}{90} \\ &\approx \underline{\underline{0.24}} \end{aligned}$$

(d) What is the probability that a randomly chosen student thought having a good social life was most important, given the student was from a suburban area?

$$\begin{aligned} P(\text{social} \mid \text{suburban}) &= \frac{22}{151} \\ &\approx \underline{\underline{0.15}} \end{aligned}$$

(e) Using conditional probability, determine whether 'Rural' and 'Grades' are independent.

$$\begin{aligned} P(\text{rural} \mid \text{grades}) &= \frac{57}{247} \\ &\approx 0.23 \end{aligned}$$

$$P(\text{ Rural }) = \frac{149}{478} = 0.31$$

Since $P(\text{rural} \mid \text{grades}) \neq P(\text{rural})$, 'Rural' and 'Grades' are not independent, i.e. dependent.