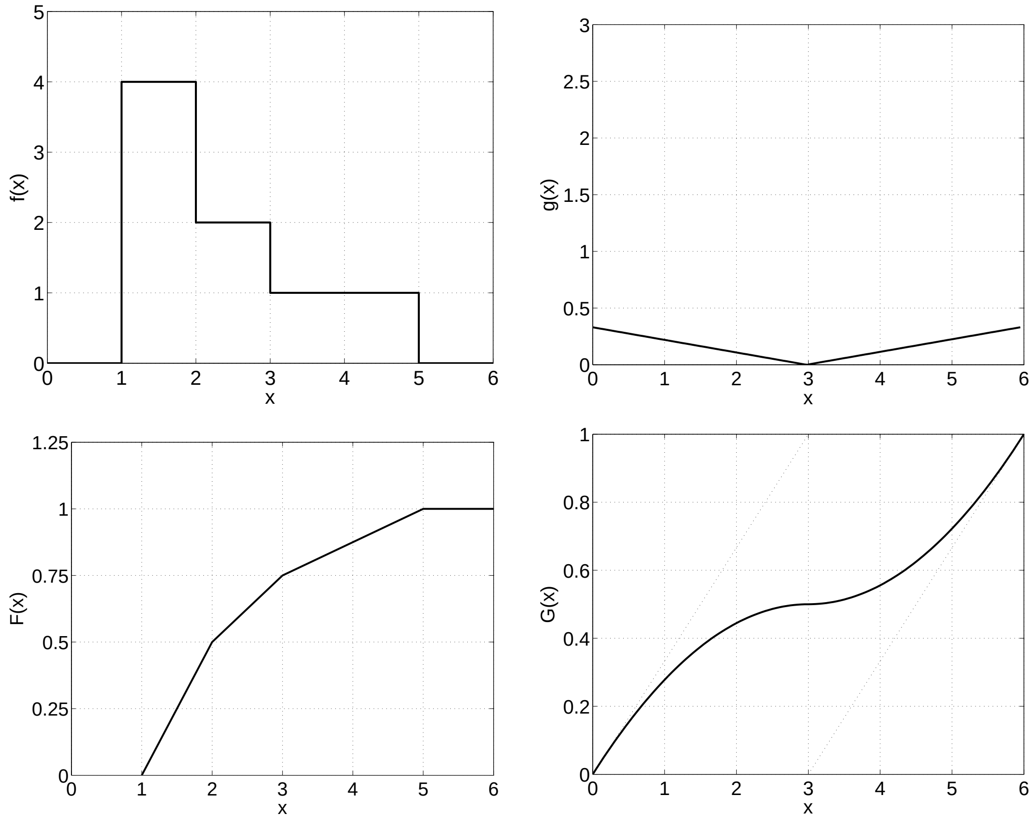
STAT 391
Final Exam Solutions

©Marina Meilă
mmp@cs.washington.edu

Problem 1 (9 points)

$Z_f = 8$ $Z_g = 1$



1.1 Compute the normalization constant Z_f for the density f in the graph on the left. 1 p.

$Z_f = 8$

1.2 Make a plot of the CDF $F(x)$ that corresponds to the density f in the graph on the left. (Be as precise as possible). 2 p.

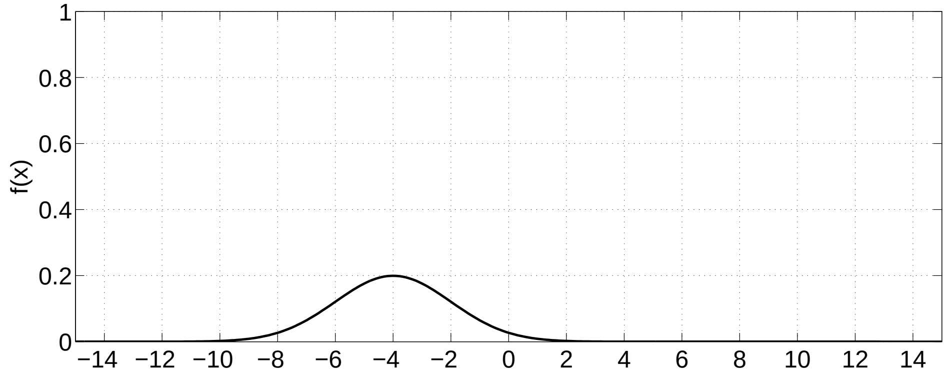
Shown above.

1.3 Make a plot of the density $g(x)$ that corresponds to the CDF G in the graph on the right. (Be as precise as possible). You can choose to draw an unnormalized density; in this case enter the normalization constant Z_g in the space above the graph. 3 p.

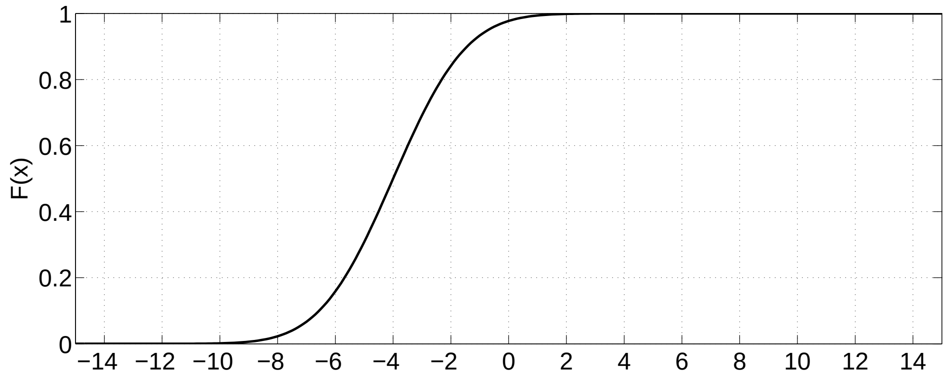
Shown above.

1.4 Make a plot of the density and CDF of the Normal distribution $N(-4,4)$. Mark the locations of μ, $\mu+\sigma, \mu-\sigma$ on the X-axis of each plot. FYI: $1/\sqrt{2\pi} \approx 0.4$, $\sqrt{2\pi} \approx 2.5$ 3 p.

$Normal(-4,4)$ density



$Normal(-4,4)$ CDF



Problem 2 (3 points)

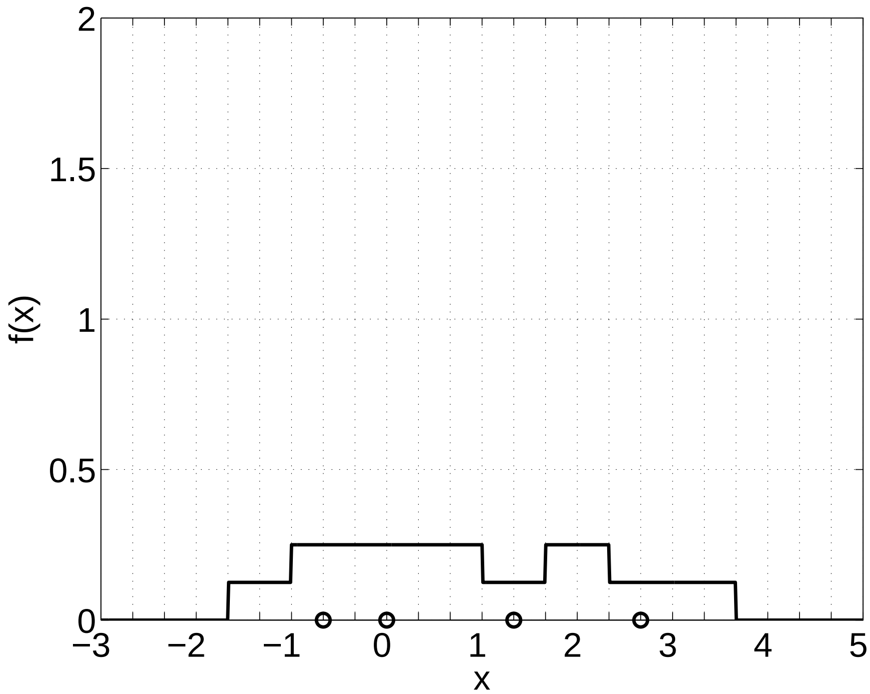
The figure depicts a data set of $n = 4$ points, $\mathcal{D} = \{-2/3, 0, 4/3, 8/3\}$, on the real axis. Draw a kernel density estimate $f(x)$ based on this data set, using a **square kernel** with $h = 2$.

You can choose to draw an unnormalized density; in this case write value of the normalization constant Z above the graph. If your density is normalized, write $Z = 1$ above the graph.

The square kernel is described by

$$k(z) = \begin{cases} 1 & \text{if } z \in [-0.5, 0.5] \\ 0 & \text{otherwise} \end{cases}$$

$Z \quad = \quad 1$



Problem 3. The Laplace density is defined on $S = (-\infty, \infty)$ by (5 points)

$$f(x) \;=\; \frac{1}{Z} e^{-\lambda|x|}$$

where $|x|$ denotes the magnitude of x, and $\lambda > 0$ is a parameter.

3.1 Calculate the value of the normalization constant Z as a function of the parameter λ. 1 p.

$$\begin{aligned}
Z &= \int_{-\infty}^{\infty} e^{-\lambda|x|} dx \\
&= 2\int_{0}^{\infty} e^{-\lambda|x|} dx \quad \text{(since } e^{-\lambda|x|} \text{ is symmetrical about 0)} \\
&= \frac{2}{-\lambda}(e^{-\infty} - e^{0}) \\
&= \frac{2}{\lambda}
\end{aligned}$$

3.2 We observe the data $\mathcal{D} = \{x_1,\, x_2,\, \ldots x_n\}$. Write the expression of the log-likelihood for this data set. 2 p.

$$
\begin{aligned}
l &= \log \prod_{i=1}^{n} \frac{1}{Z} e^{-\lambda |x_i|} \\
&= \log \prod_{i=1}^{n} \frac{\lambda}{2} e^{-\lambda |x_i|} \\
&= \log \frac{\lambda^n}{2^n} e^{-\lambda \sum_{i=1}^{n} |x_i|} \\
&= n \log \lambda - n \log 2 - \lambda \sum_{i=1}^{n} |x_i|
\end{aligned}
$$

3.3 Find the expression of the Maximum Likelihood estimate of λ based on the log-likelihood derived above. Calculate the numerical value of λ^{ML} for the dataset $\mathcal{D} = \{1, -4, 0, -3, 2\}$. 2 p.

We will maximize the log-likelihood by setting its derivative to zero.

$$
\frac{dl}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} |x_i| = 0
$$

Solving the above gives

$$
\lambda^{ML} = \frac{n}{\sum_{i=1}^{n} |x_i|}
$$

Plugging in the values for the given dataset, we get

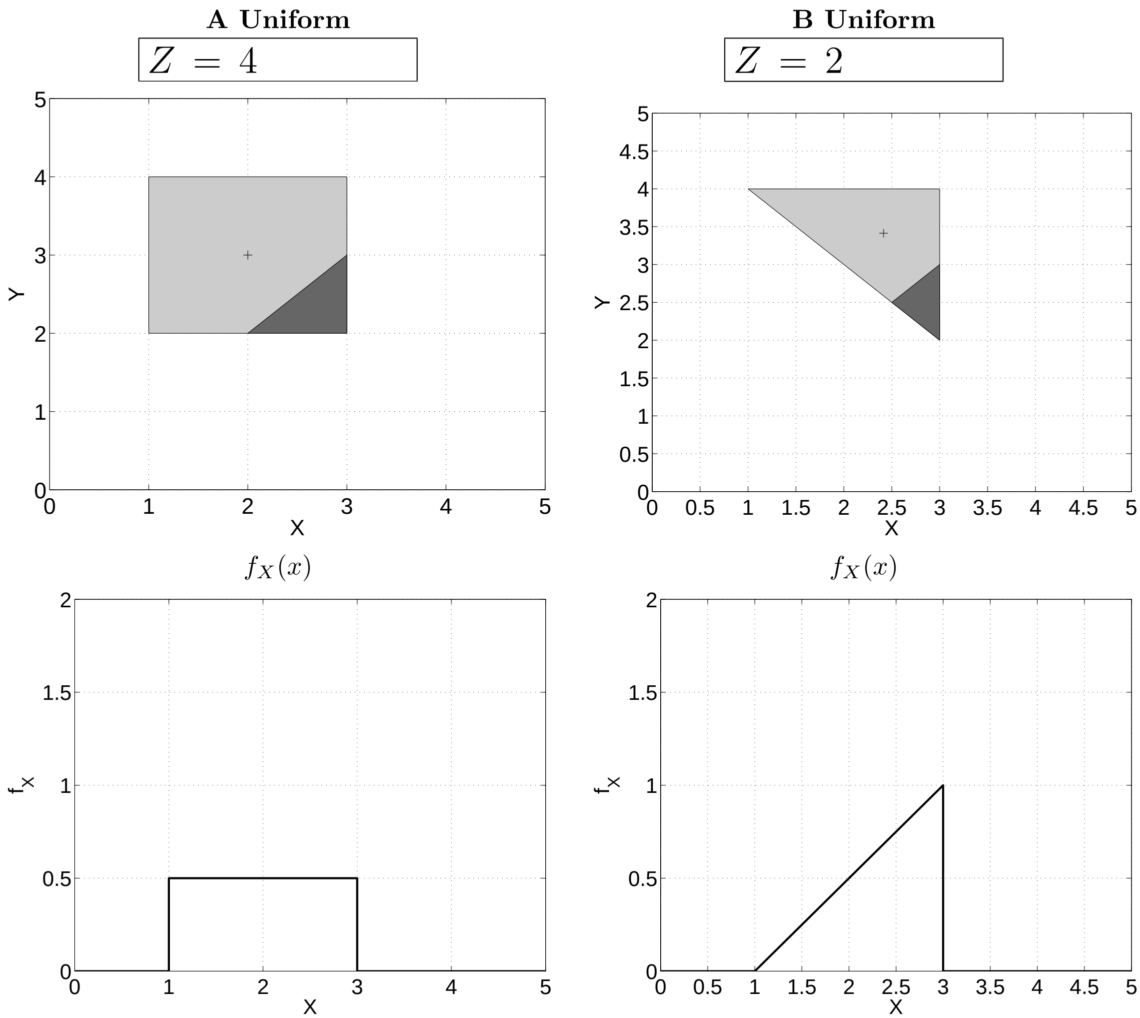
$$
\lambda^{ML} = \frac{5}{1 + 4 + 0 + 3 + 2} = 0.5
$$

Problem 4 (21 points)

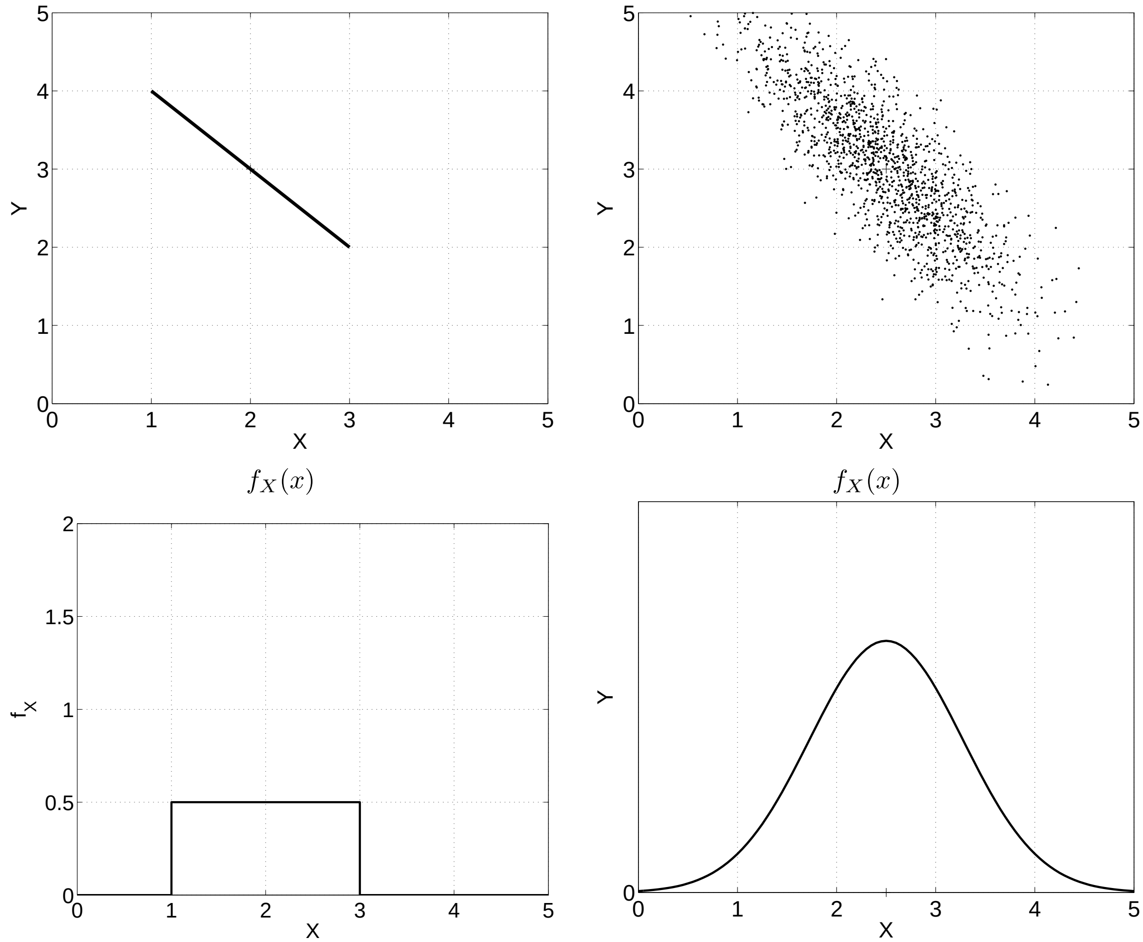
The following are bivariate densities over $(-\infty, \infty)$. The first three are unnormalized uniform densities which take value 1 over the filled in domain. The fourth one is Gaussian, and points from the distribution are shown to indicate the density.

4.1 Enter the normalization constant Z corresponding to the densities in **A, B** above their graphs. 2 p.

4.2 Plot f_X the marginal density of X below the respective joint densities. 5 p.


A Uniform
Z = 4
B Uniform
Z = 2
Y
X
f_X(x)
f_X
X
Y
X
f_X(x)
f_X
X

C Uniform **D Normal**



4.3 Are X and Y independent? 4 p.

A: Yes **B: No**
C: No **D: No**

4.4 Identify the location of the point $(E[X], E[Y])$ on each of the graphs. 5 p.

A: (2, 3) **B: $(\frac{7}{3}, \frac{10}{3})$**
C: (2, 3) **D: (2.5, 3.0)**

4.5 The correlation coefficient ρ_{XY} satisfies: 3 p.

A	C	D
$\rho_{\mathbf{XY}} = \mathbf{0}$	$\rho_{\mathbf{XY}} < \mathbf{0}$	$\rho_{\mathbf{XY}} < \mathbf{0}$

4.6 What is $P[Y \leq X]$ for **A, B**. 2 p.

A: $P[Y \leq X] = \frac{1}{8}$.

It is just $\frac{1}{Z}$ × (Area of the darkened triangle in the figure for A).

B: $P[Y \leq X] = \frac{1}{8}$.
It is just $\frac{1}{Z}$ × (Area of the darkened triangle in the figure for B).

Problem 5 (4 points)

5.1 Prove that for any two random variables X, Y 2 p.

$$Cov(X,Y) \;=\; E[XY] - E[X]E[Y]$$

For convenience, we will denote E[X], E[Y] by μ_x, μ_y respectively in our solution.

$$\begin{aligned} Cov(X,Y) &= E[(X-\mu_x)(Y-\mu_y)] \\ &= E[XY - \mu_y X - \mu_x Y + \mu_x\mu_y] \\ &= E[XY] - \mu_y E[X] - \mu_x E[Y] + \mu_x\mu_y \text{ (by linearity of expectations)} \\ &= E[XY] - \mu_y\mu_x - \mu_x\mu_y + \mu_x\mu_y \\ &= E[XY] - \mu_x\mu_y \\ &= E[XY] - E[X]E[Y] \end{aligned}$$

5.2 For two random variables X, Y, we denote $Cov(X,Y) = c$. Calculate $Cov(2X,3Y)$ as a function of c. 2 p.

Using the result proved in **5.1** we get:

$$\begin{aligned} Cov(2X,3Y) &= E[2X \cdot 3Y] - E[2X] \cdot E[3Y] \\ &= E[6XY] - 2E[X] \cdot 3E[Y] \text{ (by linearity of expectations)} \\ &= 6E[XY] - 6E[X]E[Y] \\ &= 6(E[XY] - E[X]E[Y]) \\ &= 6Cov(X,Y) \;=\; 6c. \end{aligned}$$

Problem 6 (18 points)

The CSE building has an intelligent elevator, named Ella, who carries visitors to the 6 floors of the building and to the basement (floor 0). In her idle time, Ella exercises her mind with probability questions. Can you help her find the answers?

Assume in the following that even if a person goes to floor 1, they will still stop by the elevator. Also assume that people entering the building request a floor independently of the floors chosen by other people.

The probability θ_j that a given person entering the building goes to floor j is given in the table below, which also shows the locations of some research groups in the building.

Probability θ_j	**Floor**	**Groups**
$\theta_6 = 6/40$	**6**	G S
$\theta_5 = 8/40$	**5**	T AI S CB
$\theta_4 = 5/40$	**4**	
$\theta_3 = 6/40$	**3**	AI S
$\theta_2 = 5/40$	**2**	
$\theta_1 = 7/40$	**1**	
$\theta_0 = 3/40$	**0**	G

6.1 What is the probability that a person goes to a floor higher than 3? 1 p.

Probability $= (\theta_4 + \theta_5 + \theta_6) = \frac{19}{40}$.

6.2 In front of the elevator are 4 persons. What is the probability that at least one of them is going to a floor higher than 3? (Literal answer only) 1 p.

Let X denote the event that at least one of the four is going to a floor higher than 3.

$$P(X) = 1 - P(\bar{X}) = 1 - (\theta_0 + \theta_1 + \theta_2 + \theta_3)^4$$

6.3 The Artificial Intelligence (AI), Computational Biology (CB), Graphics (G), Systems (S) and Theory (T) groups are located according to the "map" above. In other words, AI groups are on the 3-rd and 5-th floors, graphics on 0 and 6, etc. Assume that there are no other groups on floors 0, 3, 5 and 6, that a visitor to floor j will be going to only one group, and that groups on the same floor have equal probabilities of being visited. E.g a visitor to floor 6 will go to Graphics w.p 0.5 and to Systems w.p 0.5. Calculate the probability that a person goes each of the AI, G, S and T and CB groups. 4 p.

The probability of a group will be the sum of the probabilities of each of its individual "locations".

$$P(AI) = \frac{\theta_3}{2} + \frac{\theta_5}{4} = \frac{5}{40}$$
$$P(G) = \theta_0 + \frac{\theta_6}{2} = \frac{6}{40}$$
$$P(S) = \frac{\theta_3}{2} + \frac{\theta_5}{4} + \frac{\theta_6}{2} = \frac{8}{40}$$
$$P(T) = \frac{\theta_5}{4} = \frac{2}{40}$$
$$P(CB) = \frac{\theta_5}{4} = \frac{2}{40}$$

6.4 What is the probability that a person goes to the 5th floor, given that they are headed for the Theory or the Systems group? 2 p.

Let "Five" be the event that a person goes to the 5th floor. Let T and S respectively be the events that he goes to the Theory and Systems groups. The desired probability P(Five$|T \cup S$)

$$= \frac{P(T \cup S|Five)P(Five)}{P(T \cup S)} \text{ (by Bayes Rule)}$$
$$= \frac{(P(T|Five) + P(S|Five))P(Five)}{P(T) + P(S)} \text{ (since T and S are disjoint events)}$$
$$= \frac{(\frac{1}{4} + \frac{1}{4})\frac{8}{40}}{\frac{2}{40} + \frac{8}{40}} \text{ (plugging in the values)}$$
$$= \frac{4}{10} = 0.40$$

6.5 Today is a holiday, but it is also a deadline for a major AI conference. Therefore, a person going to the AI group will stay more than 8 hours (call this time "long") with probability $p = 0.7$, while a person going to one of the other groups will stay long with probability $q = 0.2$. The duration of staying depends only on the group the person is going to and is independent of anything else. 10 p.

Alice and Bob went to the same group, which is one of the aforementioned groups (call this event E_1). Alice got off before the 4-th floor (call this event E_A) and Bob stayed long (call this event E_B). Under these conditions, what is the probability that they went to the AI group?

Answer: We will denote by "wentX" the event that Alice and Bob both went to group X. We are required to compute P(wentAI$|E_A, E_B, E_1$). Applying Bayes rule, we get:

$$P(wentAI|E_A, E_B, E_1) = \frac{P(wentAI|E_1)P(E_A, E_B|wentAI, E_1)}{P(E_A, E_B|E_1)}$$
$$= \frac{P(wentAI|E_1)P(E_A, E_B|wentAI)}{\sum_{X \in All} P(wentX|E_1)P(E_A, E_B|wentX)} \quad (1)$$

In the sum in the denominator, X ranges over all groups. For any group X, $wentX \subset E_1$. So, when conditioning on $(wentX, E_1)$ we can drop E_1 as we did above.

Note that E_A and E_B are independent given $wentX$. They depend only on the group to which Alice and Bob went. Thus, the desired probability can be written as:

$$= \frac{P(wentAI|E_1)P(E_A|wentAI)P(E_B|wentAI)}{\sum_{X \in All} P(wentX|E_1)P(E_A|wentX)P(E_B|wentX)} \quad (2)$$

E_A denotes Alice getting off below the 4th floor. For a group X, P($E_A|wentX$) $= 0$ if X does not have a lab below the 4th floor. The groups T and CB do not have labs below the 4th floor. So they disappear from the denominator. The expression in (2) becomes:

$$= \frac{P(wentAI|E_1)P(E_A|wentAI)P(E_B|wentAI)}{\sum_{X \in \{AI,S,G\}} P(wentX|E_1)P(E_A|wentX)P(E_B|wentX)} \quad (3)$$

E_1 is the union of all the $wentX$ events.

$$E_1 = wentAI \cup wentS \cup wentG \cup wentT \cup wentCB$$

Recall that in problem 6.3 we computed P(X) (probability of a person going to X) for all groups X. We will now use these values for computing P($wentX|E_1$).