## ASSIGNMENT 1 PART II

**Do not forget to include a statement of academic integrity within the statement that you provide in the answer area to part II. Finally, note that you are required to show your work for full credit --- correct numeric answers may earn you little credit unless you show your work. [DUE DATE Feb 2nd]**

**QUESTION 1 [7 marks]:**

The contingency table below summarizes the proportion of job candidates by education ("high-school or less" versus "University or more") and by age (with three age categories as shown in the table). According to this table, the fraction of all those surveyed who are in the "Senior" age category and in the "high-school or less" category was 0.16. Note Young includes those 25-44 years of age, "Mid-aged" includes those falling in the age range of 45 upto (but not including) 65 years of age, and Seniors are 65 years of age or older.

a) [1.5 marks] Based on this information please complete the contingency table

| | Young | Mid-aged | Senior | Total |
|---|---|---|---|---|
| High.Sch. or less | | | 0.16 | 0.60 |
| University or more | 0.25 | | | 0.40 |
| **Total** | 0.44 | 0.34 | 0.22 | |

Assume that these study results are representative of the entire pool of job applicants for a certain company.

b) [1 marks] Is a randomly chosen applicant's educational-attainment independent of age-category? Defend your answer.

c) [1 marks] Are the event "High-school or less" and the event of being in Mid-aged category disjoint? Defend your answer.

d) [1 marks] Consider a randomly chosen applicant who falls in the "High-school or less" category. What is the probability that they are younger than 65years of age?

e) [1 marks] What is the probability that a randomly chosen applicant is either Mid-aged or has an education level of "University or more"?

f) [1 marks] You randomly select 5 candidates, what is the probability that at least 1 of these candidates is younger than 65 years of age?

g) [0.5 marks] Ignore all probabilities except the marginal probabilities, and then specify what probability value would apply to the joint event "Young" and "High.Sch or less" IF these two events were independent.

**QUESTION 2 [15 marks]:**
Some large retail stores rely on "background-check agencies" (BCA's) to help screen retail job applicants. Their aim is to screen out theft-risk employees (those who steal cash or retail-products from the store).

Reliance upon BCA's is controversial because of so-called false positives, which occur when honest, upstanding, candidates are incorrectly flagged by the BCA process as "theft risk." Assume that a background check (BC) labels an applicant as either "theft-risk" or "Not theft-risk."

A labor union has studied the problem by submitting the names of 1000 honest ("innocent") candidates to BCA's. Table 1 shows the results of these 1000 background checks:

**Table 1: BC results among the innocent.**

| **Flagged by BC as "theft risk"** | **Flagged by BC as "Not theft risk"** |
|---|---|
| 80 candidates | 920 candidates |

The BCA lobby group points to their own study which examined 25 employees convicted of work-related theft ("guilty"). Using care, they determined what the BC of each candidate would have been had it been commissioned prior to hiring (see Table 2).

**Table 2: BC results among the guilty**

| **Flagged by BC as "theft risk"** | **Flagged by BC as "Not theft risk"** |
|---|---|
| 20 candidates | 5 candidates |

Based on the results in Tables 1 and 2 (note that these are conditional distributions):

(a) [1 marks] What is the probability of being labeled "theft risk" among the innocent?

(b) [1 marks] What is the probability of being labeled "theft risk" among the guilty?

(c) [1 mark] In order to determine the probability that an applicant would end up committing theft, one might be tempted to observe that a total of 1025 employees were surveyed in the two separate studies outlined above, and to note that 100 out of those 1025 employees were "guilty" thus estimating the probability as 100/1025. This would be **incorrect**. Explain why (think of how the data was collected).

A separate national survey reveals that among the general population of retail job applicants only 1 out of every 250 will end up committing a theft in their next job ("guilty"). Use this result (when needed) for your solutions to later questions.

(d) [3.5 marks] Draw a probability tree representing the information provided for this problem, making sure to annotate the probability tree with appropriate labels and appropriate probabilities (as per example in class).

(e) [1 marks] What is the probability that a randomly chosen applicant would be labeled “theft risk” by a BC ?

(f) [3.5 marks] Construct a contingency table for this problem, being sure to specify all the joint and marginal probabilities (you may find the probability tree useful for obtaining some of these probabilities).

(g) [1.5 marks] Given that a person is labeled “theft risk” what is the probability that they are actually **innocent**?

(h) [0.5 marks] Given that a person is labeled “theft risk” what is the probability that they are actually **guilty**?

**(i)** [2 marks] Your retail company is considering a BC policy, which involves submitting all applicants to a BCA, and screening out all those labeled by the BCA as “theft risk.” Comment on the advise-ability of this policy (four sentences maximum). You may find it helpful to compute new probabilities in addition to quoting some of your answers above in order to support your recommendation.