Assignment 1 Part II

Note: Your submission will consist of two steps. First, use the answer-box to provide a statement to the TA alerting them that a PDF document containing your answers to this assignment's Part II questions has been submitted (document sharing). Second, use the shared-documents tool within MyStatLab to upload your PDF document. 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 1st]

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 ">64" age category and in the "high-school or less" category was 0.16.

a) [2 marks] Based on this information please complete the contingency table

	25-44	**45-64**	**>64**	**Total**
High.sch. or less	0.19	0.25	0.16	0.60
University or more	0.25	0.09	0.06	0.40
Total	0.44	0.34	0.22	1.00

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? Defend your answer. Independence is that one event won't alter the next or other. A randomly chosen applicant's educational-attainment is not independent of age.

c) [1 marks] Are the event "High-school or less" and the event of being in age category 25-44 disjoint? Defend your answer. Disjoint is the idea of having two events that cannot happen at the same time. Therefore the event of "high-school or less" and the event of being in the age category are not disjoint as you can be 25-44 and have high-school or less

d) [1 marks] Consider a randomly chosen applicant who falls in the "University or more" category. What is the probability that they are younger than 64 years of age?
0.25 + 0.09 = 0.34/0.4 = 0.85

e) [1 marks] What is the probability that a randomly chosen applicant is either in "45-64" or the "University or more" category?
To find the probability that an applicant is "45-64" or "University or more"; you add all applicants in "45-64" and all applicants in "University or more" then subtract applicants that are "45-64 and University or more".
$(0.34 + 0.40) – 0.09 = 0.65$

f) [1 marks] You randomly select 5 candidates, what is the probability that at least 1 of these candidates is between the ages of 25 and 64?
To find the probability that at least one of the five applicants is between 25 and 64; you subtract the probability that now of the 5 applicants is 25 to 64 from 1. Therefore you subtract the 64 or old from 1.
$1 – (0.22)^5 =$
$1 – 0.0005 = 0.9995$

QUESTION 2 [14 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"**
50 candidates	950 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:

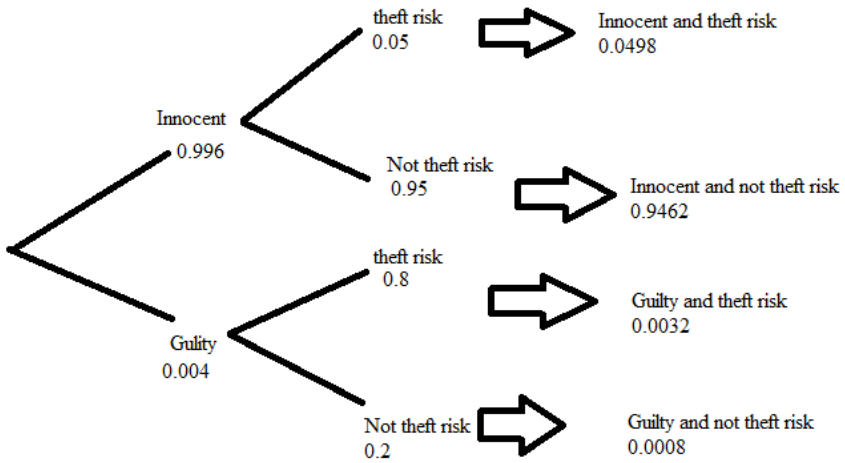
a) [1 marks] What is the probability of being labeled "theft risk" among the innocent?
The probability of being labeled "theft risk" among innocent is 50/1000. Therefore 0.05 is the probability.

b) [1 marks] What is the probability of being labeled "theft risk" among the guilty?
The probability of being labeled "theft risk" among the guilty is 20/25. Therefore 0.8 is the probability.

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 70 out of those 1025 employees were "guilty" thus estimating the probability as 70/1025. This would be **incorrect**. Explain why. The reason for using the split tables is to avoid double counting. There are only 1000 applicants therefore if you use 1025 you are double counting some applicants. You must use the compliment rule to subtract out these double applicants if you use this number.

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 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 ?

You must add up the innocent applicants and guilty applicants that are "theft risk". 0.0498 + 0.0032 = 0.053

f) [3 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).

	Theft risk	**Not theft risk**	**Total**
Innocent	0.0498	0.9462	0.996
Guilty	0.0032	0.0008	0.004
Total	0.053	0.947	**1.00**

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

P[I|Th] = P[I&Th] / P[Th]
P[I|Th] = 0.0498 / 0.053
P[I|Th] = 0.9396

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

P[G|Th] = P[G&Th] / P[Th]
P[G|Th] = 0.0032 / 0.053
P[G|Th] = 0.0604

(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.

It is not a good advisory as there is a larger number of innocent applicants than guilty applicants. Therefore when the company screens out all applicants that are “theft risks” they are eliminating a larger percentage of innocent people than guilty people who are a theft risk. Only 5.3% of applicants are a theft risk; 4.98% are innocent therefore the company loses 93.96% of theft risk that are actually innocent.