Probability Review

The following identities will be useful for these exercises:

Marginalization or "summing out". Note that Y and Z can represent sets of variables. Given a table with the full joint PD we can calculate the probability of Y by adding up all of the rows that match.	$P(Y) = \sum_{z \in Z} P(Y, z)$
Definition of conditional probability	$P(X \mid Y) = \dfrac{P(X, Y)}{P(Y)}$
Chain rule	$P(X, Y) = P(Y \mid X)P(X) = P(X \mid Y)P(Y)$
Bayes Theorem	$P(Y \mid X) = \dfrac{P(X \mid Y)P(Y)}{P(X)}$

1. **Estimating Joint Probability Distributions**

   It is your mission to develop a probabilistic model that will enable the JMU administration to uncover North Korean secret agents among the student body. The JMU secret police have randomly pulled in 20 individuals and determined their agent status as well as whether or not they golf, and what sort of hat they wear (Fedora or Fez). The data set below represents the raw data.

(loyal, golf, fedora)	(spy, golf, fedora)
(loyal, no_golf, fedora)	(loyal, golf, fedora)
(loyal, no_golf, fedora)	(spy, golf, fez)
(loyal, no_golf, fez)	(loyal, no_golf, fedora)
(loyal, no_golf, fedora)	(loyal, golf, fez)
(spy, no_golf, fedora)	(loyal, golf, fez)
(loyal, golf, fedora)	(loyal, golf, fez)
(loyal, golf, fedora)	(loyal, no_golf, fez)
(spy, golf, fez)	(loyal, no_golf, fedora)
(spy, golf, fez)	(loyal, no_golf, fedora)

   (a) Use the data above to fill in the following joint probability table.

Spy	Golfer	Fedora	Probability
T	T	T	
T	T	F	
T	F	T	
T	F	F	
F	T	T	
F	T	F	
F	F	T	
F	F	F	

(b) Do you think these probabilities are likely to be accurate (in our hypothetical world)? I.e. do you think they represent the true probability of seeing these outcomes if we draw a random student from the population? Why or why not?

(c) According to your probability distribution, what is the probability that a randomly selected student is a spy who golfs and wears a fez?
In other words, what is $P(Spy = \text{T}, Golf = \text{T}, Fedora = \text{F})$?

(d) What is the marginal probability $P(Spy = \text{T})$? Show your work.

(e) What is the marginal probability $P(Golf = \text{T}, Fedora = \text{F})$? Show your work.

(f) Use Bayes Theorem to calculate $P(Spy = \text{T} \mid Golf = \text{T}, Fedora = \text{F})$ Show your work.