Sociology 592 - Homework #1 - Descriptive and univariate statistics, intro to probability

1. In a study for discrimination between men and women in the hiring practices for a certain firm, the following starting salaries were paid to nine men and nine women hired for jobs specifying "no previous experience necessary."

Type of job	# Males hired # Females hired		Hourly rate
File clerk	2	7	\$4.28
Messenger	1	0	\$4.32
Shipping clerk	2	1	\$4.41
Assembler type I	3	1	\$4.60
Assembler type II	1	0	\$4.67

a. Find the average wage paid to men and the average wage paid to women. Do you think the difference is large enough to indicate discrimination?

b. Find the average for all 18 workers.

2. The following frequency distribution shows weekly sales of the Snoopy Hat Corporation for last year (52 weeks).

Weekly Sales	Frequency (# weeks)
\$0 - 5,999	4
\$6,000 - 11,999	6
\$12,000 - 17,999	10
\$18,000 - 23,999	16
\$24,000 - 29,999	12
\$30,000 - 35,999	4

a. Compute the mean and the mode

b. Compute the population variance

3. Prove that the following two formulas for the population variance are mathematically equivalent.

$$\frac{1}{N}\sum_{i=1}^{N}(X_{i}-\mu)^{2}=\frac{1}{N}\sum_{i=1}^{N}X_{i}^{2}-\mu^{2}$$

HINT: Keep in mind that $\Sigma X_i = n\mu$. Expand the square. That is, start with

$$\frac{1}{N}\sum_{i=1}^{N}(X_{i}^{2}-2X_{i}\mu+\mu^{2})=\frac{1}{N}\sum_{i=1}^{N}X_{i}^{2}-\mu^{2}$$

4. If A and B are mutually exclusive events, and P(A) = .20, P(B) = .40, find

 a.
 $P(A \cup B)$

 b.
 P(A | B)

 c.
 P(B | A)

5. If A and B are independent events, and P(A) = .20, P(B) = .40, find

a.	$P(A \cup B)$
b.	$P(A \mid B)$
C.	$P(B \mid A)$

6. This is a story of three prisoners. Al, Bernie, and Chuck are condemned to death, but at the last minute the governor has decided to pardon one of them. He has told the prison warden but on the condition that he is not to reveal who is to go free. The warden likes to bring good news, but true to his word, all he can do is tell them that one of them will go free. They plead for more details, but he has made a promise.

Al has an idea: "I'll speak to the warden privately. One of you, of course, is going to die, so I'll ask him for a name. He won't tell the person involved, but maybe he'll tell me."

Alone with Al in his office, the warden agrees and tells Al that Chuck is one of the ones who is going to die. "Aha," says Al, "I've just increased my chance of survival from 1/3 to 1/2."

Indicate whether you or not you agree with Al, and why. If you disagree, state what you think the chances of survival are for Al and Bernie.

7. On August 11, 1980, Time Magazine presented a profile of the unemployed in America. Use the chart to answer the following questions:

a. According to the figure, what is the probability that a randomly selected person who is unemployed will be either adult or white, i.e. $P(W \cup A)$? What is $P(W \cap A)$?

b. Given that an unemployed person is an adult, what is the probability that the person is white? What is P (W)?

c. Is the event W independent of event A? Explain, using your answer to part (b).

d. Find the value of P (A) P (W). Use this value and the results from part (a) to determine whether A and W are independent.



Source: Time Magazine, August 11, 1980

8. A study on the probability that a randomly selected person smokes cigarettes divided the US population into three age groups: Under 30 (<30), between 30 and 50 (30-50), and over 50 (>50). Half those under 30 were found to smoke.

a. If P(<30) = 1/2, find the probability that a randomly selected person is under 30 and smokes.

b. If P (Smokes | <30) = 1/2, P (Smokes | 30-50) = 1/4, and P (Smokes | >50) = 1/2, does this indicate independence or dependence between age and smoking?

c. If P (30-50) = 1/4 and P (>50) = 1/4, find P (Smokes).

d. Replace the probability symbols in the following table with their appropriate values.

	<30	30-50	>50	
Smokes (S)	$P(S \cap <30)$	$P(S \cap 30-50)$	$P(S \cap >50)$	P (S)
Doesn't (\overline{S})	$P(\overline{S} \cap <30)$	P ($\overline{S} \cap 30-50$)	$P(\overline{S} \cap >50)$	$P(\overline{S})$
	P (<30)	P (30-50)	P (>50)	

e. Find P (Smokes $\cup > 50$).

9. [OPTIONAL] Use SPSS and/or Stata to confirm part or all of your answers to problems 1 and 2. Remember, the variance estimates will differ because the programs assume you are analyzing a sample rather than a population.