More on scatterplots

Example 1: Does fast driving waste fuel? How fast does the fuel consumption of a car change as its speed increases? Today’s Minitab worksheet contains the data for a British Ford Escort. Speed is measured in kilometers per hour, and fuel consumption is measured in liters of gasoline used per 100 kilometers traveled.
(a) Make a scatterplot. Which is the explanatory variable?
(b) Describe the form of the relationship. Why is it not linear? Explain why the form of the relationship makes sense.

(c) It does not make sense to describe the variables as either positively associated or negatively associated. Why?

(d) Is the relationship reasonably strong or quite weak? Explain your answer.

Chapter 4 Probability and Probability Distributions

Randomness

• "Random" does not mean "haphazard"
• Randomness is a kind of order that emerges only in the long run.
Example 2:
(a) A fair coin is to be flipped once. Can you predict the outcome?
(b) A fair coin is to be tossed 100 times. Can you predict what proportion of the flips will be heads? Perform an experiment to see if your prediction is correct.

Bernoulli Trials
Some experiments have only two possible outcomes, for example, flipping a (possibly unfair) coin, or asking a randomly selected person a "yes" or "no" question. A trial of such experiments (such as flipping the coin 100 times) is called a Bernoulli Trial, or Bernoulli experiment. The individual outcomes of the trials are designated as either "successes" or "failures". For example, we could call heads a "success" in the coin tossing experiment.

Example 3: Simulating an opinion poll. A recent opinion poll showed that about 73% of married women agree that their husbands do at least their fair share of household chores. Suppose that this is exactly true. Choosing a married woman at random then has probability 0.73 of getting one who agrees that her husband does his share.

(a) Using Minitab, simulate drawing 20 women, then 80 women, then 320 women. What proportion agree in each case? We expect (but because of chance variation we can’t be sure) that the proportion will be closer to 0.73 in the longer run of trials.

(b) Simulate drawing 20 women 10 times and record the percents in each trial who agree. Then simulate drawing 320 women 10 times and again record the ten percents. Which set of 10 results is less variable? We expect the results of 320 trials to be more predictable (less variable) than the results of 20 trials. That is "long-run regularity" showing itself.

Experiments
An experiment is a random phenomenon, such as
- Tossing a pair of coins
- Dealing a poker hand
- Selecting 10 light bulbs at random and finding their average life span

Probability models
- The sample space $S$ of an experiment is the set of all possible outcomes of the experiment.
- An event is any outcome or a set of outcomes of an experiment. That is, an event is a subset of the sample space.
- A probability model is a sample space $S$ and a way of assigning probabilities to events in $S$. 

2
**Example 2:** A manager is to select two of Bob, Tom and Mary to go on a business trip.

(a) What is the sample space?

(b) What is the probability that Bob is chosen?

**Probability definitions**

Let $P(A)$ denote the probability of the event $A$.

- The *complement* $\overline{A}$ of the event $A$ is the event that $A$ does *not* occur.

- The *union* $A \cup B$ of the events $A$ and $B$ is the event that either $A$ or $B$ occurs.

- The *intersection* $A \cap B$ of the events $A$ and $B$ is the event that both of $A$ and $B$ occurs.

- The events $A$ and $B$ are mutually exclusive if the event $A \cap B$ is impossible.

**Probability rules**

1. $0 \leq P(A) \leq 1$.
2. If $S$ is the sample space, then $P(S) = 1$.
3a. If $A$ and $B$ are mutually exclusive, then
   
   \[ P(A \cup B) = P(A) + P(B). \]

3b. In general,
   
   \[ P(A \cup B) = P(A) + P(B) - P(A \cap B). \]

4. $P(\overline{A}) = 1 - P(A)$. 
