MATH 1010 Vocabulary for Chapter 11: Probability

Phenomenon: an occurrence or possible occurrence. In statistics, is classified as 1 of 2 types: 1) deterministic: something that can be predicted exactly based on existing information. 2) Random: cannot be predicted exactly due to a lack of information (usually about events in the future, such as weather forecasts).

Experiment: any observation or measurement of a random phenomenon.

Outcomes: the possible results of an experiment, the total number of events that could happen as a result of the phenomenon occurring. Example: the phenomenon is a front moves in; the outcomes are changes in temperature, humidity, cloudiness, or precipitation.

Sample space (S): the set of all possible results. n(S) = the total number of possible results or the number of observations in the experiment. {Think: the table of 36 possible outcomes for rolling a pair of dice}

Event: generally, something that happens or could happen; an occurrence. Examples are some phenomenon such as an experiment, such as a coin flip. A dice roll, a spinner stopping on a space, a meteorite landing in Murfreesboro, snow, rain, answering a question correctly, traffic signal turning red as you drive toward the intersection, Dr. O giving a quiz, or anything similar.

Event (E): in statistics, specifically, a subset of the sample space; the particular collection of possible outcomes in which we are interested. n(E) = the number of favorable outcomes.

NOTE: $n(E) \leq n(S)$. {the number of favorable outcomes is less than or equal to the total number of possible outcomes}

Favorable outcomes: outcomes that belong to the event, (a.k.a. successes). {Think: rolling a pair of dice and getting a double}

Probability: the likelihood of an Event happening: the number of outcomes in the Event, **n(E)**, divided by the total number of outcomes in the sample space, **n(S)**.

P(E) = n(E)/n(S). The probability that an outcome in the Event will occur is a fraction of the total possible outcomes and will **ALWAYS** be a number between 0 and 1 inclusive: $0 \leq P(E) \leq 1$. P(E) may be written as a rational fraction, a decimal, or a percent.

Examples:

1) What is the probability of a double when rolling a pair of dice (event is rolling a double)?

n(E) = 6, the six outcomes where both die have the same number showing

n(S) = 36, the total number of outcomes possible when rolling a pair of dice

P(E) = 6/36 = ⅙ or approximately 0.1667

2) What is the probability that the sum of the numbers on the pair of dice is between 1 and 13?
n(E) = 36, all outcomes have a sum or 2, 3, 4, ..., or 12
n(S) = 36, the total number of outcomes possible when rolling a pair of dice

P(E) = 36/36 = 1

3) What is the probability that the sum of the numbers on the pair of dice is 1 or 13?
n(E) = 0, all outcomes have a sum or 2, 3, 4, ..., or 12
n(S) = 36, the total number of outcomes possible when rolling a pair of dice

P(E) = 0/36 = 0

Theoretical probability: mathematically determined (coin toss, card drawn, or roll of die or dice).

Empirical probability: experimentally determined (data are given as the results of a survey, questionnaire, observations, trials).

Odds: a method of expressing the likelihood of an event as a comparison of favorable outcomes to unfavorable outcomes (successes to failures). The sum of the two numbers MUST equal the total possible outcomes. When flipping a fair coin, the probability of getting a head is ½ (1 favorable outcome out of 2 possible outcomes), but the odds of a head are 1:1 (read 1 to 1; 1 favorable outcome compared to 1 unfavorable outcome).

Odds in favor: if all outcomes in a sample space are equally likely, some number, ***a***, of them are favorable to Event E, and the remaining number of outcomes, ***b***, are unfavorable, then the odds in favor of E are ***a:b (a to b)***.
Example: What are the odds in favor of a double when rolling a pair of dice (event is rolling a double)?
a = 6, the six outcomes where both die have the same number showing
b = 30, the number of outcomes not showing doubles when rolling a pair of dice
The odds in favor of rolling a double are 6:30, simplified to 1:5 1 way of rolling a double compared to 5 ways of rolling something else.

Odds against: when the situation is the same as above, the odds against E are ***b:a*** (***b*** to ***a***).
Example: What are the odds against a double when rolling a pair of dice (event is rolling a double)?
a = 6, the six outcomes where both die have the same number showing
b = 30, the number of outcomes not showing doubles when rolling a pair of dice

The odds against rolling a double are 30:6, simplified to 5:1 5 ways of rolling something else compared to 1 way of rolling a double.

Conditional probability of B given A, P(B|A): the probability of Event B happening computed on the assumption that Event A has already happened.

Independent events: two events are independent if the knowledge that one event happened does NOT change the probability that the other event can happen. If Event A happens, the probability the Event B happens does NOT change.

Probability of "**not** E": If E is an event, then the probability of "not E" MUST be 1 – P(E). Example: The probability of rolling a 2 on a fair die is 1/6, the probability of not rolling a 2 is 5/6. The two probabilities MUST add to exactly 1.

Multiplication Rule of Probability: If A and B are any two events, then the probability of both occurring is calculated as follows:

If A and B are dependent, then P(A and B) = P(A) • P(B|A)

Example: When 2 cards are drawn without replacement from a standard deck, what is the probability that the second card is a face card given that the first card was a queen?

P(A) = 1/13; P(B) = 11/51

If A and B are independent, then P(A and B) = P(A) • P(B)

Example: When 2 cards are drawn with replacement from a standard deck, what is the probability that the second card is a face card given that the first card was a queen?

P(A) = 1/13; P(B) = 3/13

MATH 1010 Vocabulary for Chapter 12: Statistics

Statistics: gathering, organizing, and analyzing data. Usually of two types:

a) descriptive statistics: collecting, organizing, summarizing, and presenting data that describes someone or something. Examples: sex (M/F), ethnicity, age, weight, height, hair color, etc.
b) inferential statistics: using data gathered and analyzed from a sample to draw inferences or conclusions about a population.

Data: information (factual) about an item in a population. A single datum from a data set is symbolized as x. Three main types:

a) raw data: collected but unprocessed information, not evaluated yet. Example: You got 9 correct, but you do not know yet whether that is a good score or not;
b) quantitative data: an objective measure (a number), usually how many of a particular item; Example: the section has 22 students, 12 female and 10 male.

c) qualitative data: a subjective measure that describes a characteristic of an item; generally not a number. Example: the first section is inquisitive while the second section is thoughtful.

Population: the total number of items of interest. Example: the batch of 8000 radios.

Sample: a subset of the population that includes some, but generally not all, of the items in the population. Sometimes referred to as the data set. Example: the sample of 14 radios.

n: the number of items under consideration and may be a population n or a sample n.

Frequency distribution: a list of the distinct data values (each x) along with their frequency (f). A list of each made score and how many made that score.

Visual displays of data: graphs and charts to give a visual presentation of data and to serve as tools of exploratory data analysis. Main types:

a) Histogram: connected bars that show the frequency as the height and the data score (x) or class along the bottom.
b) Frequency polygon: a graph with the frequency plotted as a single point and the series of points connected by line segments that connect to the horizontal axis on each end; a type of line graph.
c) Stem-and-leaf display: a type of frequency visual where the stem is the tens digits listed vertically in the left column and the leaves are the ones digits listed horizontally to the right. Allows the researcher to see the frequency and the actual scores within a data set.
d) Bar graph: similar to a histogram but with gaps between the bars.
e) Circle graph (aka pie chart): a circle divided into segments (slices of the pie) where the size of the segment shows the relative magnitude of the data category at some point in time.
f) Line graph: connected data points that shows change over a period of time, shows trends in the data.

Measure of Central Tendency: a middle value that is representative of a data set. Three main types:

a) mean: arithmetic average: the sum of all scores divided by the number of scores.
b) x Most commonly used. Symbol: x-bar, an x with a line above it. An example of a weighted mean is a GPA.
c) median: the exact value that divides the data set into two groups with exactly the same number of data points in each. The median may or may not be one of the data points. To find the mean: first rank and count the data points; next, if an odd number of data points the median is the middle item in the list; if an even number, it is the average of the 2 middle numbers.
d) mode: the most frequently occurring item or items in a data set. Not every data set has a mode, some sets have one, some sets have two or more (a set with exactly two is called bimodal)

Dispersion: the spread of a data set; how tightly or loosely the data gathers around the middle.

Range: a measure of dispersion calculated by: highest value – lowest value.

Standard Deviation: one of the most useful measures of dispersion; based on the deviation of scores from the mean (score – mean). Scores smaller than the mean have a negative deviation; score higher than the mean have a positive deviation. The sum of deviations for a data set is ALWAYS zero!