

The Normal Curve

Ray Block, Jr.
PS 585 (Research Methods)
Fall, 2003

Distributions

Last Class

- Probability:
 - The Logic Behind It
 - [Some] Rules to It

Today's Class

- The Normal Curve
 - Frequency Distributions
 - Probability Distributions
 - The Normal Distribution

Frequency Distributions

Frequency Distributions

- What are they?
- A method of organizing and presenting data
- Take the following example (to right): Frequencies for 10 Rolls of a Fair Die

Frequency Distributions

- Example: shows the number of times each observation occurs when the values are arranged in order of magnitude

Frequency Distributions

- What's the purpose?
- Raw data provide little, if any useful information...

Frequency Distributions

- ...Frequency distributions convert raw data into useful information

Frequency Distributions

- Frequency distributions can also tell us:
 - How the observations cluster around a central value
 - The degree of difference between observations

Frequency Distributions

- For example:
 - Rolling a 3 was the most frequently occurring observation
 - The possible observations range from rolling a 1 to rolling a 6

Probability Distributions

Probability Distributions

- What are they?
- A listing of all the values the random variable can assume with their corresponding probabilities

Probability Distributions

- Example (to right): The probability distribution that results from the rolling of a single fair die.

Probability Distributions

- They are just like a frequency distribution, except that is based on theory, not data

Probability Distributions

- It is a frequency distribution of a random variable if an experiment were conducted a large number of times

The Normal Curve

The Normal Curve

- What is it?
- A special type of *probability distribution* dealing with a special type of *random variable*

The Normal Curve

- Random variable:
 - A variable whose outcome is uncertain until a random “experiment” is done

The Normal Curve

- Two type of random variables:
 - Discrete = Countable # of outcomes
 - Example: A coin flip
 - Continuous = Assume any value w/in a given range
 - Example: A book can weigh from 0.5 lbs to 10 lbs

The Normal Curve

- The normal curve is probably one of the most important and widely used *continuous* distribution

The Normal Curve

- Sometimes referred to as:
 - The *Standard Normal* distribution
 - The distribution of a *normal random* variable

The Normal Curve

- A Discrete (Binomial) Example:
 - What is the probability of me flipping a coin and it landing heads if I flipped it x number of times?
 - If x is a really large number, then I cannot actually perform this “experiment”
 - But I can infer what the distribution would look like based on my understanding of probability distributions

The Normal Curve

- What’s the purpose?
- If you add enough observations, even a binomial distribution eventually turns into a normal-looking distribution

The Normal Curve

- The central limit theorem:
 - Implies that the larger the sample size, the closer the probability distribution will resemble to a normal distribution

The Normal Curve

- This tells us that the normal distribution is the “mother of all distributions” (Gonick and Smith 1993, 88)

The Normal Curve

- The binomial distribution occurs commonly in nature
- It is not hard to understand
- However, it is tough to calculate

The Normal Curve

- Normal distributions are less intuitive, but easier to calculate
- Therefore, we can use the normal to approximate binomial distributions

The Normal Curve

- Characteristics of the Normal Distribution:
 - It is bell shaped and is symmetrical about its mean
 - It is asymptotic to the axis, i.e., it extends indefinitely in either direction from the mean
 - It is a continuous distribution.
 - It is a family of curves i.e., every unique pair of mean and standard deviation defines a different normal distribution. Thus, the normal distribution is completely described by two parameters: mean and standard deviation.
 - Total area under the curve sums to 1 (the area of the distribution on each side of the mean is 0.5).
 - It is uni-modal (values mound up only in the center of the curve)
 - The probability that a random variable will have a value between any two points is equal to the area under the curve between those points

References

- FYI:
 - Levin, Jack and James Alan Fox. 2003. Elementary Statistics In Social Research, 9th Edition. Boston, MA: Pearson Education Group, Inc.
 - Wonnacott, Thomas H. and Ronald J. Wonnacott. 1990. Introductory Statistics, 5th Edition. New York, NY: John Wiley & Sons.