Introduction to Statistics (Econometrics)

Descriptive Statistics

Panagiotis Th. Konstantinou

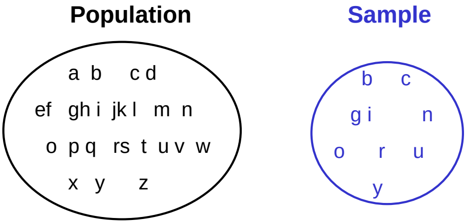
MSc in International European Economics Studies/Economics and Law in Energy Markets,

Athens University of Economics and Business

First Draft: July 15, 2015. **This Draft**: September 13, 2025.

Key Concepts

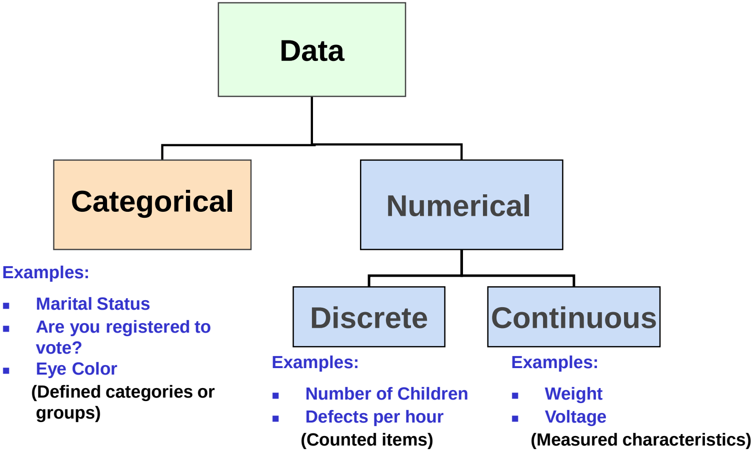
- A **population** is the collection of all items of interest or under investigation (N represents the population size)
- A **sample** is an observed subset of the population (n represents the sample size)
- A **parameter** is a specific characteristic of a population
- A **statistic** is a specific characteristic of a sample



Values calculated using population data are called parameters

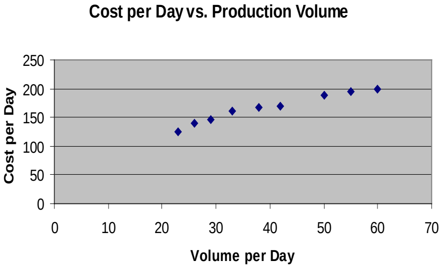
Values computed from sample data are called statistics

Data Types


Data
Categorical
Numerical
Examples:
Marital Status
Are you registered to vote?
Eye Color
(Defined categories or groups)
Discrete
Continuous
Examples:
Number of Children
Defects per hour
(Counted items)
Examples:
Weight
Voltage
(Measured characteristics)

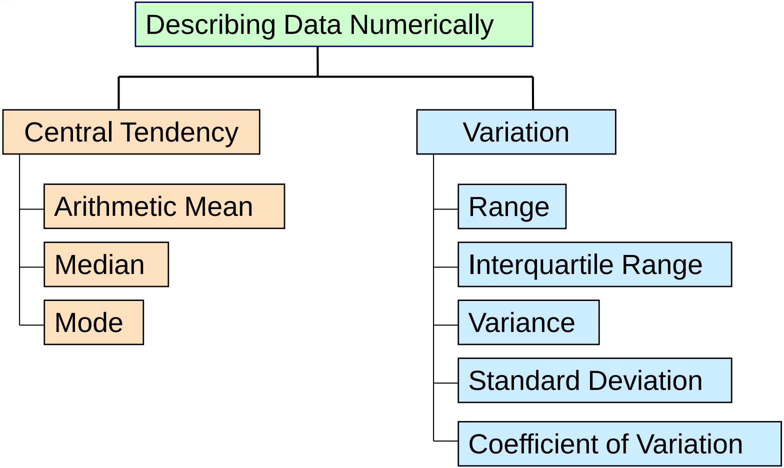
Relationships Between Variables

Volume per day	Cost per day
23	125
26	140
29	146
33	160
38	167
42	170
50	188
55	195
60	200

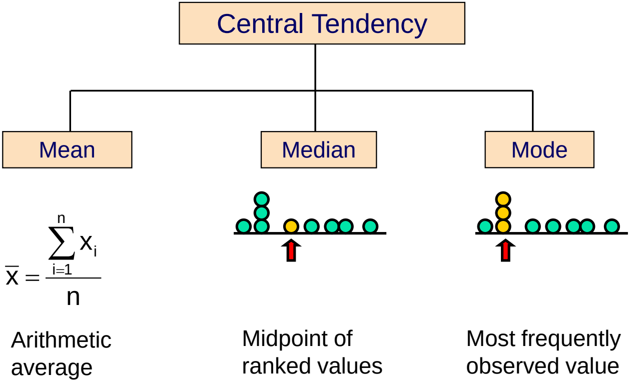


Investment Category	Investor A	Investor B	Investor C	Total
Stocks	46.5	55	27.5	**129**
Bonds	32.0	44	19.0	**95**
CD	15.5	20	13.5	**49**
Savings	16.0	28	7.0	**51**
Total	**110.0**	**147**	**67.0**	**324**

Describing Data Numerically



Measures of Central Tendency



- Median position $\frac{n+1}{2}$ position in the ordered data
  - If the number of values is odd, the median is the middle number
  - If the number of values is even, the median is the average of the two middle numbers

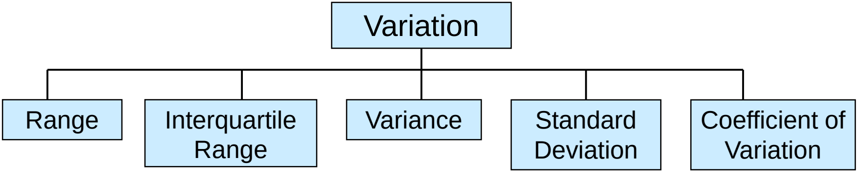
Measures of Central Tendency

Example

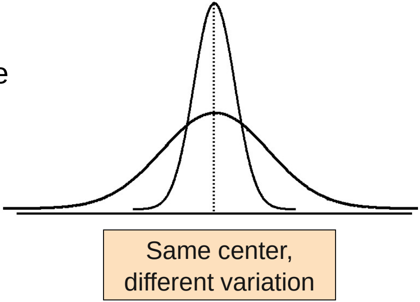
House Prices	
	\$2,000,000
	500,000
	300,000
	100,000
	100,000
Sum	\$3,000,000

- **Mean**: \$3,000,000/5 = \$600,000
- **Median**: middle value of ranked data = \$300,000
- **Mode**: most frequent value = \$100,000

Measures of Variability



- Measures of variation give information on the spread or variability of the data values.



Variance

- ***Population Variance***: Average of squared deviations of values from the mean

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

where

  - μ = population mean
  - N = population size
  - $X_i = i-$th value of the variable X

- ***Sample Variance***: Average (approximately) of squared deviations of values from the sample mean:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

where

  - $\bar{x}$ = sample mean/average
  - n = sample size
  - $x_i = i-$th value of the variable X

Standard Deviation

- ***Population Standard Deviation***: Most commonly used measure of variation
  - Shows variation about the mean
  - Has the *same units as the original data*

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

- ***Sample Standard Deviation***: Most commonly used measure of variation
  - Shows variation about the *sample* mean
  - Has the *same units as the original data*

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

Standard Deviation

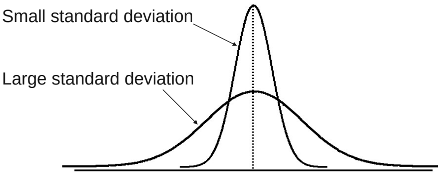
Example: Sample Standard Deviation Computation

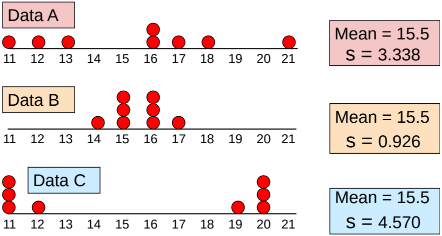
- Sample Data (x_i) : 10 12 14 15 17 18 18 24
- $n = 8$ and sample mean $= \bar{x} = 16$
- So the standard deviation is

$$
\begin{aligned}
s &= \sqrt{\frac{(10-\bar{x})^2 + (12-\bar{x})^2 + (14-\bar{x})^2 + \cdots + (24-\bar{x})^2}{n-1}} \\
&= \sqrt{\frac{(10-16)^2 + (12-16)^2 + (14-16)^2 + \cdots + (24-16)^2}{8-1}} \\
&= \sqrt{\frac{126}{7}} = 4.2426
\end{aligned}
$$

- This is a measure of the “**average**” scatter around the (sample) mean.

Comparing Standard Deviations





- The smaller the standard deviation, the more concentrated are the values around the mean.

- Same mean, different standard deviations.

Coefficient of Variation

- Measures relative variation and is always in percentage (%)
- Shows variation **relative to mean**
- Can be used to compare two or more sets of data **measured in different units**

$$CV = \left(\frac{s_x}{\bar{x}}\right) \cdot 100\%$$

- **Stock A**:
  - Avg price last year = \$50
  - Standard deviation = \$5

$$CV_A = \left(\frac{\$5}{\$50}\right) \cdot 100\% = 10\%$$

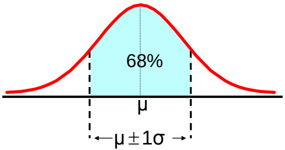
- **Stock B**:
  - Avg. price last year = \$100
  - Standard deviation = \$5

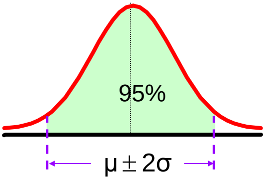
$$CV_B = \left(\frac{\$5}{\$100}\right) \cdot 100\% = 5\%$$

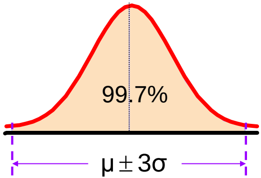
- Both stocks have the same standard deviation, but stock B is less variable relative to its price

The Empirical Rule

If the data distribution is bell-shaped, then the interval:







- $\mu \pm 1\sigma$ contains about 68% of the values in the population or the sample
- $\mu \pm 2\sigma$ contains about 95% of the values in the population or the sample
- $\mu \pm 3\sigma$ contains almost all (about 99.7%) of the values in the population or the sample.

Covariance

- The covariance measures the strength of the linear relationship between **two variables**
- The ***population covariance***:

$$\mathrm{Cov}(X, Y) = \sigma_{XY} = \frac{\sum_{i=1}^{N}(X_i - \mu_X)(Y_i - \mu_Y)}{N}.$$

- The ***sample covariance***:

$$\widehat{\mathrm{Cov}(x, y)} = s_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n - 1}.$$

- Only concerned with the strength of the relationship
- No causal effect is implied
  - $\mathrm{Cov}(x, y) > 0$, x and y tend to move in the ***same*** direction
  - $\mathrm{Cov}(x, y) < 0$, x and y tend to move in ***opposite*** directions

Correlation Coefficients

- The correlation coefficient measures the relative strength of the linear relationship between **two variables**
- The ***population correlation coefficient***:

$$\text{Corr}(X, Y) = \rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

- The ***sample correlation coefficient***:

$$\widehat{\text{Corr}(x, y)} = r_{xy} = \frac{\widehat{\text{Cov}(x, y)}}{s_x s_y}.$$

- Unit free and ranges between -1 and 1
  - The closer to -1, the stronger the negative linear relationship
  - The closer to 1, the stronger the positive linear relationship
  - The closer to 0, the weaker any positive linear relationship

Introduction to Statistics (Econometrics)

Elements of Probability Theory

Panagiotis Th. Konstantinou

MSc in International European Economics Studies/Economics and Law in Energy Markets,

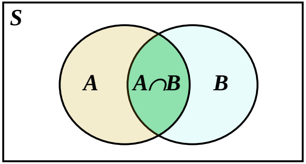
Athens University of Economics and Business

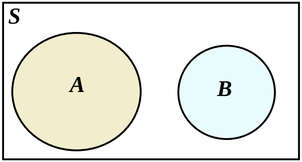
First Draft: July 15, 2015. **This Draft**: September 14, 2025.

Important Terms in Probability – I

- ***Random Experiment*** – it is a process leading to an uncertain outcome
- ***Basic Outcome*** (S_i) – a possible outcome (the most basic one) of a random experiment
- ***Sample Space*** (S) – the collection of all possible (basic) outcomes of a random experiment
- ***Event*** A – is any subset of basic outcomes from the sample space ($A \subseteq S$). This is our object of interest here – among other things.

Important Terms in Probability – II

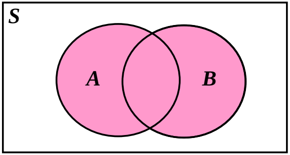


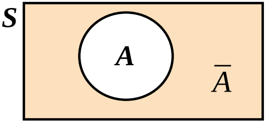


- ***Intersection of Events*** – If A and B are two events in a sample space S, then their intersection, $A \cap B$, is the set of all outcomes in S that belong to **both** A and B

- We say that A and B are ***Mutually Exclusive Events*** if they have no basic outcomes in common i.e., the set $A \cap B$ is empty ($\varnothing$)

Important Terms in Probability – III





- ***Union of Events*** – If A and B are two events in a sample space S, then their union, $A \cup B$, is the set of all outcomes in S that belong to either A or B

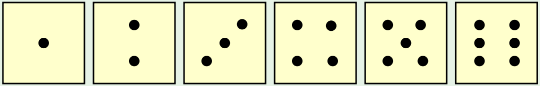
- The ***Complement*** of an event A is the set of all basic outcomes in the sample space that do not belong to A. The complement is denoted $\bar{A}$ or A^c.

Important Terms in Probability – IV

- Events $E_1, E_2, \ldots, E_k$ are ***Collectively Exhaustive*** events if $E_1 \cup E_2 \cup \ldots \cup E_k = S$, i.e., the events completely cover the sample space.

Examples

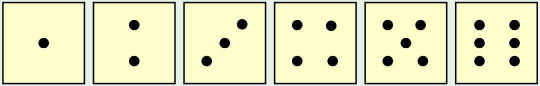
Let the ***Sample Space*** be the collection of all possible outcomes of rolling one die $S = \{1, 2, 3, 4, 5, 6\}$.



- Let ***A*** be the event "Number rolled is even": $A = \{2, 4, 6\}$
- Let ***B*** be the event "Number rolled is at least 4" : $B = \{4, 5, 6\}$
- ***Mutually exclusive***: A and B are **not** mutually exclusive. The outcomes 4 and 6 are common to both.

Important Terms in Probability – V

Examples (Continued)



$A = \{2, 4, 6\} \qquad B = \{4, 5, 6\}$

- ***Collectively exhaustive***: A and B are **not** collectively exhaustive. $A \cup B$ does not contain 1 or 3.
- ***Complements***: $\bar{A} = \{1, 3, 5\}$ and $\bar{B} = \{1, 2, 3\}$
- ***Intersections***: $A \cap B = \{4, 6\}$; $\bar{A} \cap B = \{5\}$; $A \cap \bar{B} = \{2\}$; $\bar{A} \cap \bar{B} = \{1, 3\}$.
- ***Unions***: $A \cup B = \{2, 4, 5, 6\}$; $A \cup \bar{A} = \{1, 2, 3, 4, 5, 6\} = S$.

Assessing Probability – I

- ***Probability*** **–** the chance that an uncertain event A will occur is always between 0 and 1.

$$\underbrace{0}_{\text{Impossible}} \leq \Pr(A) \leq \underbrace{1}_{\text{Certain}}$$

- There are three approaches to assessing the probability of an uncertain event:

Assessing Probability – II

1. ***Classical Definition of Probability***:

$$\text{Probability of an event } A = \frac{N_A}{N}$$

$$= \frac{\text{number of outcomes that satisfy the event } A}{\text{total number of outcomes in the sample space } S}$$

- Assumes all outcomes in the sample space are equally likely to occur.
- ***Example***: Consider the experiment of tossing 2 coins. The sample space is $S = \{HH, HT, TH, TT\}$.
- Event $A = \{\text{one } T\} = \{TH, HT\}$. Hence $\Pr(A) = 0.5$ – assuming that all basic outcomes are equally likely.
- Event $B = \{\text{at least one } T\} = \{TH, HT, TT\}$. So $\Pr(B) = 0.75$.

Assessing Probability – III

2. ***Probability as Relative Frequency***:

$$\text{Probability of an event } A = \frac{n_A}{n}$$

$$= \frac{\text{number of events in the population that satisfy event } A}{\text{total number of events in the population}}$$

- The limit of the proportion of times that an event A occurs in a large number of trials, n.

Assessing Probability – IV

3. ***Subjective Probability***: an individual has opinion or belief about the probability of occurrence of A.
    - When economic conditions or a company's circumstances change rapidly, it might be inappropriate to assign probabilities based solely on historical data
    - We can use any data available as well as our experience and intuition, but ultimately a probability value should express our degree of belief that the experimental outcome will occur.

Measuring Outcomes – I

Classical Definition of Probability

- ***Basic Rule of Counting***: If an experiment consists of a sequence of k steps in which there are n_1 possible results for the first step, n_2 possible results for the second step, and so on, then the total number of experimental outcomes is given by $(n_1)(n_2)...(n_k)$ – tree diagram...

Measuring Outcomes – II

Classical Definition of Probability

- ***Counting Rule for Combinations*** (Number of Combinations of n Objects taken k at a time): A second useful counting rule enables us to count the number of experimental outcomes when k objects are to be selected from a set of n objects (the ordering does not matter)

$$C_k^n = \binom{n}{k} = \frac{n!}{k!(n-k)!},$$

where $n! = n(n-1)(n-2)...(2)(1)$ and $0! = 1$.

Measuring Outcomes – III

Classical Definition of Probability

- ***Example***: Suppose we flip three coins. How many are the possible combinations with (exactly) 1 T?

$$C_1^3 = \binom{3}{1} = \frac{3!}{1!(3-1)!} = 3.$$

- ***Example***: Suppose we flip three coins. How many are the possible combinations with *at least* $1T$?
- ***Example***: Suppose that there are two groups of questions. Group A with 6 questions and group B with 4 questions. How many are the possible half-a-dozens we can put together?

$$n = 6 + 4 = 10; \; C_6^{10} = \binom{10}{6} = \frac{10!}{6!(10-6)!} = 210.$$

Measuring Outcomes – IV

Classical Definition of Probability

- ***Example***: How many possible half-a-dozens we can put together, preserving the ratio 4 : 2?

$$\binom{6}{4} \times \binom{4}{2} = 15 \times 6 = 90.$$

- **Probability**: What is the probability of selecting a particular half-a-dozen (with ratio 4 : 2), when we choose at random? Using the classical definition of probability

$$\frac{90}{210} = 0.4286$$

Measuring Outcomes – V

Classical Definition of Probability

- ***Counting Rule for Permutations*** (Number of Permutations of n Objects taken k at a time): A third useful counting rule enables us to count the number of experimental outcomes when k objects are to be selected from a set of n objects, **where the order of selection is important**

$$P_k^n = \frac{n!}{(n-k)!}.$$

Measuring Outcomes – VI

Classical Definition of Probability

- ***Example***: How many 3-digit lock combinations can we make from the numbers 1, 2, 3, and 4?
  The order of the choice is important! So

$$P_3^4 = \frac{4!}{1!} = 4! = 4(3)(2)(1) = 24.$$

- ***Example***: Let the characters A, B, Γ. In how many ways can we combine them in making triads?

$$P_3^3 = \frac{3!}{0!} = 3! = 3(2)(1) = 6.$$

  These are: $AB\Gamma, A\Gamma B, BA\Gamma, B\Gamma A, \Gamma AB$,and ΓBA.

Measuring Outcomes – VII

Classical Definition of Probability

- ***Example***: Let the characters A, B, Γ, Δ, E. In how many ways is it possible to combine them into pairs?

∗ If the order matters, we may have

$$P_2^5 = \frac{5!}{3!} = (5)(4) = 20.$$

∗ If the order does not matters, we may choose pairs

$$C_2^5 = \binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{5!}{2!3!} = 10$$

Probability Axioms

- The following ***Axioms*** hold

1. If A is any event in the sample space S, then

$$0 \leq \Pr(A) \leq 1.$$

2. Let A be an event in S, and let S_i denote the basic outcomes. Then

$$\Pr(A) = \sum_{\text{all } S_i \text{ in } A} \Pr(S_i).$$

3. $\Pr(S) = 1.$

Probability Rules – I

- The ***Complement Rule***:

$$\Pr(\bar{A}) = 1 - \Pr(A) \text{ [i.e., } \Pr(A) + \Pr(\bar{A}) = 1].$$

- The ***Addition Rule***: The probability of the union of two events is

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

- Probabilities and joint probabilities for two events A and B are summarized in the following table:

	B	$\bar{B}$	
A	$\Pr(A \cap B)$	$\Pr(A \cap \bar{B})$	$\Pr(A)$
$\bar{A}$	$\Pr(\bar{A} \cap B)$	$\Pr(\bar{A} \cap \bar{B})$	$\Pr(\bar{A})$
	$\Pr(B)$	$\Pr(\bar{B})$	$\Pr(S) = 1$

Probability Rules – II

Example (Addition Rule)

Consider a standard deck of 52 cards, with four suits ♡♣♢♠. Let event A = card is an Ace and event B = card is from a red suit.

Pr(Red ∪ Ace) = Pr(Red) + Pr(Ace) - Pr(Red ∩ Ace)

= 26/52 + 4/52 - 2/52 = 28/52

Type	Color		Total
	Red	Black	
Ace	2	2	4
Non-Ace	24	24	48
Total	26	26	52

Don't count the two red aces twice!

Conditional Probability – I

- A ***conditional probability*** is the probability of one event, given that another event has occurred:

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} \text{ (if } \Pr(B) > 0);$$

$$\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)} \text{ (if } \Pr(A) > 0)$$

Conditional Probability – II

Example (Conditional Probability)

Of the cars on a used car lot, 70% have air conditioning (AC) and 40% have a CD player (CD). 20% of the cars have both. What is the probability that a car has a CD player, given that it has AC?
$[\Pr(CD|AC) = ?]$

	CD	No CD	Total
AC	.2	.5	.7
No AC	.2	.1	.3
Total	.4	.6	1.0

$$\Pr(\mathrm{CD}|\mathrm{AC}) = \frac{\Pr(\mathrm{CD} \cap \mathrm{AC})}{\Pr(\mathrm{AC})} = \frac{.2}{.7} = .2857$$

Multiplication Rule

- The ***Multiplication Rule*** for two events A and B:

$$\Pr(A \cap B) = \Pr(A|B)\Pr(B) = \Pr(B|A)\Pr(A)$$

Example (Multiplication Rule)

Pr(Red ∩ Ace) = Pr(Red| Ace)Pr(Ace)

$$= \left(\frac{2}{4}\right)\left(\frac{4}{52}\right) = \frac{2}{52}$$

$$= \frac{\text{number of cards that are red and ace}}{\text{total number of cards}} = \frac{2}{52}$$

Type	Color		Total
	Red	Black	
Ace	2	2	4
Non-Ace	24	24	48
Total	26	26	52

Statistical Independence – I

- Two events are ***statistically independent*** if and only if:

$$\Pr(A \cap B) = \Pr(A)\Pr(B).$$

  - Events A and B are independent when the probability of one event is not affected by the other event.
  - If A and B are independent, then

$$\begin{aligned} \Pr(A|B) &= \Pr(A), \text{ if } \Pr(B) > 0; \\ \Pr(B|A) &= \Pr(B), \text{ if } \Pr(A) > 0. \end{aligned}$$

Statistical Independence – II

Example (Statistical Independence)

Of the cars on a used car lot, 70% have air conditioning (AC) and 40% have a CD player (CD). 20% of the cars have both. Are the events AC and CD statistically independent?

	CD	No CD	Total
AC	.2	.5	.7
No AC	.2	.1	.3
Total	.4	.6	1.0

$P(AC \cap CD) = 0.2$

$P(AC) = 0.7$
$P(CD) = 0.4$
} **$P(AC)P(CD) = (0.7)(0.4) = 0.28$**

$P(AC \cap CD) = 0.2 \neq P(AC)P(CD) = 0.28$

So the two events are not statistically independent

Statistical Independence – III

Remark (Exclussive Events and Statistical Independence)

Let two events A and B with $\Pr(A) > 0$ *and* $\Pr(B) > 0$ *which are mutually exclusive. Are A and B independent?* ***NO!***

To see this use a Venn diagram and the formula of conditional probability (or the multiplication rule).

- If one mutually exclusive event is known to occur, the other cannot occur; thus, the probability of the other event occurring is reduced to zero (and they are therefore dependent).

Examples – I

- **Example 1.** In a certain population, 10% of the people can be classified as being high risk for a heart attack. Three people are randomly selected from this population. What is the probability that exactly one of the three are high risk?

- Define H: high risk, and N: not high risk. Then

$$\Pr(\text{exactly one high risk}) = \Pr(HNN) + \Pr(NHN) + \Pr(NNH) =$$
$$= \Pr(H)\Pr(N)\Pr(N) + \Pr(N)\Pr(H)\Pr(N) + \Pr(N)\Pr(N)\Pr(H)$$
$$= (.1)(.9)(.9) + (.9)(.1)(.9) + (.9)(.9)(.1) = 3(.1)(.9)^2 = .243$$

Examples – II

- **Example 2.** Suppose we have additional information in the previous example. We know that only 49% of the population are female. Also, of the female patients, 8% are high risk. A single person is selected at random. What is the probability that it is a high risk female?

- Define H: high risk, and F: female. From the example, $\Pr(F) =$ .49 and $\Pr(H|F) = .08$. Using the Multiplication Rule:

$$\Pr(\text{high risk female}) = \Pr(H \cap F)$$
$$= \Pr(F)\Pr(H|F) = .49(.08) = .0392$$

Correlation Coefficients

Examples