# Conditional Independence

## 1000: The Idea Informally

Informally, if two properties are independent, the information that someone (or something) has one of the properties provides no information about whether that someone (or something) has the other property. Two properties are conditionally independent on a third property if, **among** individuals with the third property, the first two are independent.

We write that properties A and B are independent as: A _||_ B, and that A and B are conditionally independent on C as A _||_ B | C. The idea is that, once we learn that an individual has property C, learning that they have property B tells us nothing **extra** about whether they have property A.

In frequency terms, properties A and B are independent if the frequency of A is the same as the frequency of A given B. A and B are conditionally independent on C if the frequency of A given C is the same as the frequency of A given C & B.

| | | |
|---|---|---|
| **Independence** | **A _||_ B** | Fr(**A**) = Fr(**A** \| **B**) |
| **Conditional Independence** | **A _||_ B \| C** | Fr(**A** \| **C**) = Fr(**A** \| **C**, **B**) |

FIGURE 1000-1

Consider a few examples.

### Example 1

Lets assume that the property "is bald" and the property "has a bald brother" are dependent. That is, the frequency of being bald is not the same as the frequency of being bald given you have a brother who is bald. Learning that someone "has a bald brother" is informative about whether he is bald. Now consider the property -- "has a mother with the baldness gene." Conditional on having a mother with the baldness gene, however, the properties being bald and having a bald brother are independent. Once we learn that an individual (say Frank) has a mother with the baldness gene, telling us that Frank has a bald brother tells us nothing extra about whether or not Frank is bald. That is, the frequency of being bald is the same as the frequency of being bald conditional on having a brother being bald, given that the mother has the baldness gene.

### Example 2

Consider 3-year-olds and chicken-pox. The properties: "has chicken-pox symptoms" and "was exposed to another person with chicken-pox in the past week" are dependent. Among 3-year-olds, the frequency of having chicken-pox symptoms is much lower than the frequency of having chicken-pox symptoms given exposure to another person with chicken-pox in the past week. Consider a third property: "infected with the chicken-pox virus." **Conditional** on infection, symptoms and exposure are **independent**. That is, among three year olds infected with the chicken-pox virus, the frequency of symptoms is the same as the frequency of symptoms given you were exposed to another person with chicken-pox. Put another way, once we know a 3-year-old is infected with the virus, telling us that he or she was exposed to another person with chicken pox adds no information relevant to his or her symptoms.

### Example 3

Among American High School students, college plans and gender are associated. More precisely, the properties "plans to go college" and "is male" are dependent. The frequency of HS students with college plans is not the same as the frequency of HS students with college plans among just males. Among those who have parents that strongly encourage college, however, "college plans" and "male" are **independent**. That is, among students with strong parental encouragement, the frequency of college planners is the same as the frequency of college planners among males.

Like independence, conditional independence is symmetric. That is, if A and B are independent conditional on C, then B and A are independent conditional on C.

**THE SYMMETRY OF CONDITIONAL INDEPENDENCE**

A _||_ B | C ⟷ B _||_ A | C

$$Fr(A \mid C) = Fr(A \mid C, B) \longleftrightarrow Fr(B \mid C) = Fr(B \mid C, A)$$

FIGURE 1000-2

Two (discrete) **variables** X and Y are conditionally independent on a third variable Z if, for each value of Z, each value of X is independent of each value of Y.

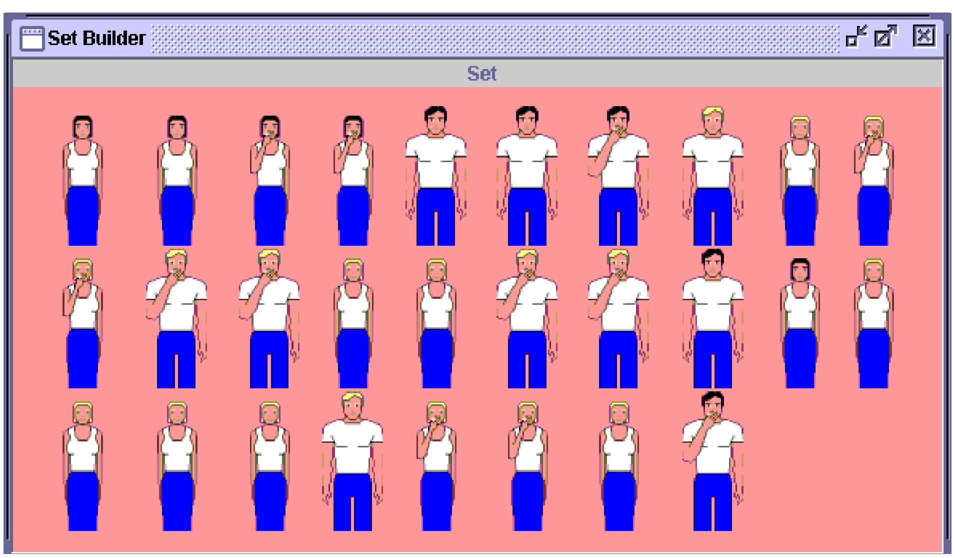
**< A link to exercises in the interactive version of this module. >**

## 2000: Conditional Independence Among Properties

## 2100: Representing Conditional Independence Among Properties

Two properties A and B are independent conditional on C if the frequency of A given C is the same as the frequency of A conditional on B & C. So if we want to use pie charts and histograms to graphically represent conditional independence, we need to display at least two charts: one for the frequency of A given C, and another for the frequency of A given B & C.

For example, suppose that we created a sample with Setbuilder, and wanted to use histograms to represent that being female was independent of smoking, given that you are dark-haired. Then we would need the following histograms:



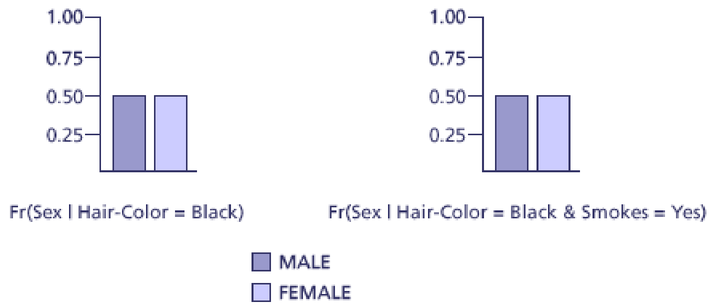



FIGURE 2100-1

**< A link to exercises in the interactive version of this module. >**

---

## 2200: Determining Conditional Independence Among Properties

---

Sometimes, instead of histograms, we are given raw data in the form of a data table or a contingency table. To determine whether two properties are conditionally independent given this kind of information, we need to apply the definition of conditional independence. Recall how the definitions of independence and conditional independence are similar.

| | | |
|---|---|---|
| **Independence** | **A** _||_ **B** | Fr(**A**) = Fr(**A** \| **B**) |
| **Conditional Independence** | **A** _||_ **B** \| **C** | Fr(**A** \| **C**) = Fr(**A** \| **C**, **B**) |

FIGURE 2200-1

Consider the following data table for 12 High School students.

**TABLE 2200-1: DATA TABLE**

| Individual | Class Year | Taken Biology | Sex |
|---|---|---|---|
| 1 | Sophomore | No | Female |
| 2 | Junior | No | Male |
| 3 | Senior | Yes | Male |
| 4 | Senior | Yes | Male |
| 5 | Junior | Yes | Male |
| 6 | Sophomore | No | Female |
| 7 | Senior | Yes | Male |
| 8 | Sophomore | No | Male |
| 9 | Sophomore | No | Female |
| 10 | Junior | No | Female |
| 11 | Senior | Yes | Female |
| 12 | Junior | Yes | Female |

Do male students take Biology more than females? Is being male independent of taking Biology? To answer this question, we need to ask whether:

$Fr_s$(Bio = Yes) = $Fr_s$(Bio = Yes | Sex = Male ).

---

## 3000: Case Studies

---

### 3100: Gender, Parental Encouragement, and College Plans

---

In 1968, Sewell and Shaw collected data on five variables from a sample of 10,318 Wisconsin high school seniors. The variables and their values are:

**TABLE 3100-1: SEWELL AND SHAW STUDY**

| | |
|---|---|
| **SEX** | [Male, Female] |
| **IQ** | [Low, Medium, High, Highest] |
| **COLLEGE PLANS** | [Yes, No] |
| **PARENTAL ENCOURAGEMENT** | [Low, High] |
| **SOCIO-ECONOMIC STATUS** | [Low, Low-medium, Medium, Medium-high, High] |

We will concentrate on three of their variables: **SEX**, **COLLEGE PLANS (CP)**, and **PARENTAL ENCOURAGEMENT (PE)**. The frequencies for these variables are as follows:

**TABLE 3100-2: GENDER, COLLEGE PLANS, AND PARENTAL ENCOURAGEMENT**

| | |
|---|---|
| Fr(**SEX** = Male) = 0.484 | Fr(**SEX** = Female) = 0.516 |
| Fr(**CP** = Yes) = 0.327 | Fr(**CP** = No) = 0.673 |
| Fr(**PE** = High) = 0.481 | Fr(**PE** = Low) = 0.519 |

Parental encouragement was widely believed to be a major influence on a student's college plans, and Sewell and Shaw's data confirmed this:

**TABLE 3100-3: THE INFLUENCE OF PARENTAL ENCOURAGEMENT**

Fr(**CP** = Yes | **PE** = Low) = 0.063
Fr(**CP** = Yes | **PE** = High) = 0.572

It was also widely believed that a student's gender was associated with his or her plans to go to college. Again, Sewell and Shaw's data confirmed this:

**TABLE 3100-4: THE INFLUENCE OF GENDER**

Fr(**CP** = Yes | **SEX** = Female) = 0.306
Fr(**CP** = Yes | **SEX** = Male) = 0.357

What was not known, however, was whether males and females got different levels of parental encouragement, and whether these differences were enough to explain the differences in college plans between the sexes. It turned out that males received much more parental encouragement than females:

**TABLE 3100-5: PARENTAL ENCOURAGEMENT AND GENDER**

Fr(**PE** = High | **SEX** = Female) = 0.464
Fr(**PE** = High | **SEX** = Male) = 0.578

The question of whether these differences were enough to explain the differences in **COLLEGE PLANS** between the sexes is a question of conditional independence. If, among students who had high **PARENTAL ENCOURAGEMENT**, **SEX** was independent of **COLLEGE PLANS**, then **SEX** and **COLLEGE PLANS** are independent conditional on **PARENTAL ENCOURAGEMENT** and there is evidence that a social policy aimed at equalizing **PARENTAL ENCOURAGEMENT** among the sexes might indeed be effective at equalizing **COLLEGE PLANS** among the sexes.

**< A link to exercises in the interactive version of this module. >**

According to Sewall and Shaw's data, the conditional independence holds quite closely, although not perfectly:

**TABLE 3100-6: THE RESULTS OF THE SEWALL AND SHAW STUDY**

Fr(**CP** = High | **PE** = High) = 0.572
Fr(**CP** = High | **PE** = High, **SEX** = Male) = 0.563
Fr(**CP** = High | **PE** = High, **SEX** = Female) = 0.581

Among those with high **PARENTAL ENCOURAGEMENT**, the frequency of **COLLEGE PLANS** is nearly the same as the frequency of **COLLEGE PLANS** given **SEX** = male, and its also nearly the same as the frequency of **COLLEGE PLANS** given **SEX** = female. In fact, even though parents are more likely to encourage males to go to college, given that they do encourage a child, females are slightly more likely to plan to go to college than are males (0.581 to 0.563).

---

## 4000: The Idea Formally

---

## 4100: Table of Notations

---

Let S be a sample or population, i.e., any non-empty collection of objects. Objects in the collection may have various properties. If A is a property, the set of objects that have A in S will be signified by "A" itself, and the set of objects that do not have A in S -- the complement of A in S -- will be signified by "~A". The set of objects that have both properties A and B is denoted by "A & B"; the set of objects that have either A or B or both is denoted by "A v B".

| | |
|---|---|
| **A** | The subset of the sample that has property **A** |
| **~A** | The subset of the sample that does not have property **A** |
| **A & B** | The subset of the sample that has both property **A** and property **B** |
| **A v B** | The subset of the sample that has either property **A**, or property **B**, or both |

FIGURE 4100-1

The **cardinality** of a set is the number of elements in the set. If S is a set, we write the cardinality of S as: |S|. For example, if I use the letter P to represent the set of major planets in our solar system, then |P| will be 9. If A represents the set of individuals in a study who were HIV positive and 5% of the 1000 people studied were HIV positive, then |A| would be 50.

Two properties A and B are said to be **exclusive** if no one in the sample has both A and B, i.e., if the subset of the sample that has both property A and property B is the empty set (A & B = ∅) . For example the properties male and female are exclusive, but the properties male and being a smoker are not.

Two properties are said to be **exhaustive** if everyone in the sample S has at least one of them (A v B = S). For example, the properties male and female are exhaustive, but the properties male and smoker are not.

---

## 4200: Definitions

### 4210: Relative Frequency

If S is a finite sample and A is a property in the sample, the relative frequency of A in S is defined to be:

$$Fr_S(\mathbf{A}) = \frac{|\,\mathbf{A}\,|}{|\,\mathbf{S}\,|}$$

FIGURE 4210-1

### 4220: Conditional Relative Frequency

If S is a sample, and A and B are properties in the sample, the frequency of A conditional on B is defined to be:

$$Fr_S(\mathbf{A} \mid \mathbf{B}) = \frac{Fr_S(\mathbf{A}\ \&\ \mathbf{B})}{Fr_S(\mathbf{B})}$$

FIGURE 4220-1

Applying the definition of frequency to both the numerator and denominator of the right hand side, we get:

$$Fr_S(\mathbf{A} \mid \mathbf{B}) = \frac{\frac{|\,\mathbf{A}\ \&\ \mathbf{B}\,|}{|\,\mathbf{S}\,|}}{\frac{|\,\mathbf{B}\,|}{|\,\mathbf{S}\,|}}$$

FIGURE 4220-2

And since the # in S cancels, we end up with:

$$Fr_S(\mathbf{A} \mid \mathbf{B}) = \frac{|\,\mathbf{A}\ \&\ \mathbf{B}\,|}{|\,\mathbf{B}\,|}$$

FIGURE 4220-3

### 4230: Independence and Association

Properties A and B are independent in sample S (which we write as A $\_||\__S$ B) if and only if the frequency of A in S equals the frequency of A conditional on B in S. Formally:

$$A \_||\__S B$$

if and only if

$$Fr_S(A) = Fr_S(A \mid B)$$

FIGURE 4230-1

Properties A and B are associated (dependent) in sample S (written as A $\not\perp\!\!\!\perp_S$ B) if and only if they are not independent in S.

An equivalent definition of independence is: Properties A and B are independent in a sample S if and only if the frequency of both A and B equals the frequency of A multiplied by the frequency of B. Formally:

$$A \_||\__S B$$

if and only if

$$Fr_S(A \,\&\, B) = Fr_S(A) * Fr_S(Q)$$

FIGURE 4230-2

This equivalent definition follows straightforwardly from the first definition of independence, and the definition of conditional relative frequency: A _||_ B if and only if Fr(A) = Fr(A | B), and Fr(A | B) = Fr(A & B) / Fr(B) by definition.

---

### 4240: Definition of Conditional Independence for Properties

---

We write: A _||_$_S$ B | C to mean A is independent of B conditional on C in sample S. Where context permits, we leave out the S. We write Fr(A, B) to mean the Fr(A & B).

Provided that Fr(C) does not equal 0 and Fr(C & B) does not equal 0, properties A and B are conditionally independent given property C in sample S if and only if the conditional frequency of A given C in S equals the conditional frequency of A given C & B in S:

$$A \_||\__S B \mid C$$

if and only if

$$Fr_S(A \mid C) = Fr_S(A \mid C, B)$$

FIGURE 4240-1

This definition follows the intuitions we have been developing, but is unnecessarily restrictive -- it requires that Fr(C & B) and Fr(C) not equal zero. This is because Fr(A | C,B) = Fr(A & C & B)/Fr(C & B), which is undefined when Fr(C & B) = 0, and Fr(A | C) = Fr(A & C) / Fr(C), which is undefined when Fr(C) = 0. A more general definition, which does not require that Fr(C & B) = 0, but is equivalent to the one above when Fr(C & B) = 0, is:

$$A \_||\__S B \mid C$$

if and only if

$$Fr_S(A \mid C) * Fr_S(B \mid C) = Fr_S(A, B \mid C)$$

FIGURE 4240-2

This definition still requires that Fr(C) not equal 0, but it is in fact the definition used most commonly.

---

### 4300: Facts about Conditional Independence

---

### 4310: Conditional Independence is Symmetric

Conditional independence is symmetric.

**CLAIM:**

**A _||_ B | C <--> B _||_ A | C**

**PROOF:**

| | | |
|---|---|---|
| 1 | **A _||_ B \| C** | Assumption |
| 2 | $Fr_S$(**A \| C**) = $Fr_S$(**A \| C, B**) | 1, Definition of Conditional Independence |
| 3 | $Fr_S$(**A, C**) / $Fr_S$(**C**) = $Fr_S$(**A, C, B**) / $Fr_S$(**C, B**) | 2, Definition of Conditional Frequency |
| 4 | $Fr_S$(**C, B**) / $Fr_S$(**C**) = $Fr_S$(**A, C, B**) / $Fr_S$(**A, C**) | 3, Algebra |
| 5 | $Fr_S$(**B, C**) / $Fr_S$(**C**) = $Fr_S$(**A, C, B**) / $Fr_S$(**C, A**) | 4, $Fr_S$(**A, B**) = $Fr_S$(**B, A**) |
| 6 | $Fr_S$(**B \| C**) = $Fr_S$(**B \| C, A**) | 5, Definition of Conditional Frequency |
| 7 | **B _||_ A \| C** | 6, Definition of Conditional Independence |

FIGURE 4310-1

## 4320: Conditional Independence of Properties and their Complements

Recall that if A is a property, then ~A refers to the property of not having A (so ~A is the complement of A). If we have two properties, A and B, that are independent conditional on a third property, C, then conditional on C, each property A (and its complement ~A) is also independent of both the other property B and its complement ~B:

**A _||_$_S$ B | C**

if and only if

**~A _||_$_S$ B | C**

if and only if

**A _||_$_S$ ~B | C**

if and only if

**~A _||_$_S$ ~B | C**

FIGURE 4320-1

For example, consider the properties Male, Smoker, and Blond. If we know that Smoker is independent of Male conditional on Blond, then we can infer that:

+ Smoker is independent of Female (the complement of Male) conditional on Blond;
+ Non-smoker (the complement of Smoker) is independent of Male conditional on Blond; and
+ Non-smoker (the complement of Smoker) is independent of Female (the complement of Male) conditional on Blond.

However, if we look within a sub-population in which every individual has the property C, and we find that A and B are independent, that doesn't mean that A and B are independent within the compliment of that sub-population, that is, one in which no individual has the property C.

**A _||_ B | C**

does **not** mean

**A _||_ B | ~C**

FIGURE 4320-2

For example, if we look within the sub-population of females, we will find that being bald and having a bald maternal grandfather are **independent** properties. If we look within the sub-population of males, however, we will find that that being bald and having a bald maternal grandfather are **dependent**

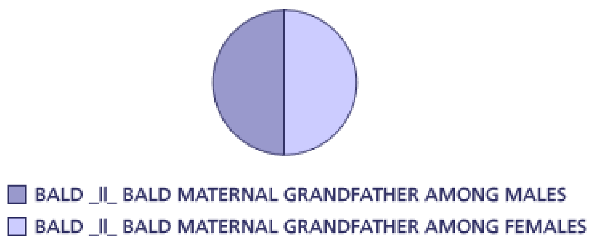



FIGURE 4320-3

---

## 5000: Conditional Independence for Variables

---

### 5100: Definition of Conditional Independence for Variables

---

Two (discrete) **variables** X and Y are conditionally independent on a third variable Z if, for **each** value of Z, **each** value of X is independent of **each** value of Y.

If X has k values, and Y has n values, and Z has m values, then the number of conditional independence relations among these values that must hold for conditional independence among the variables to hold is: (k-1) * (n-1) * (m).

For example, suppose that X is the variable INCOME, with values [low, medium, high], Y the variable EDUCATION, with values [HS, College, Graduate Degree], and Z the variable SEX, with values [male, female]. Then if INCOME and EDUCATION were independent conditional on SEX, all of the following would hold:

+ INCOME = Low _||_ EDUCATION = HS | SEX = Male
+ INCOME = Low _||_ EDUCATION = College | SEX = Male
+ INCOME = Low _||_ EDUCATION = Graduate Degree | SEX = Male
+ INCOME = Medium _||_ EDUCATION = HS | SEX = Male
+ INCOME = Medium _||_ EDUCATION = College | SEX = Male
+ INCOME = Medium _||_ EDUCATION = Graduate Degree | SEX = Male
+ INCOME = High _||_ EDUCATION = HS | SEX = Male
+ INCOME = High _||_ EDUCATION = College | SEX = Male
+ INCOME = High _||_ EDUCATION = Graduate Degree | SEX = Male
+ INCOME = Low _||_ EDUCATION = HS | SEX = Female
+ INCOME = Low _||_ EDUCATION = College | SEX = Female

+ **INCOME** = Low _||_ **EDUCATION** = Graduate Degree | **SEX** = Female
+ **INCOME** = Medium _||_ **EDUCATION** = HS | **SEX** = Female
+ **INCOME** = Medium _||_ **EDUCATION** = College | **SEX** = Female
+ **INCOME** = Medium _||_ **EDUCATION** = Graduate Degree | **SEX** = Female
+ **INCOME** = High _||_ **EDUCATION** = HS | **SEX** = Female
+ **INCOME** = High _||_ **EDUCATION** = College | **SEX** = Female
+ **INCOME** = High _||_ **EDUCATION** = Graduate Degree | **SEX** = Female

If any of these 12 independencies fails to hold (i.e., if any of the values are associated), then the two variables are associated conditional on **SEX**. The variables **INCOME** and **EDUCATION** are independent conditional on **SEX** only if, for each value of **SEX**, **every** value of **INCOME** is independent of every value of **EDUCATION**.

---

## 5200: Conditional Independence for Binary Variables

---

Recall that a binary variable has only two values. Since the values of variables are just properties, this means that the values of a binary variable are a property and its complement. We earlier saw that if a property A is independent of another property B, conditional on a third property C, then the complement of the first, ~A is also independent of B conditional on C.

So, if one value of a binary variable is independent of the value of another variable conditional on a value of a third, then the other value of the first binary variable must also be independent of that value of the second variable conditional on that value of the third. And this means that, if one value of a binary variable is independent of a second value (property) or variable conditional on a value of a third variable, then the first binary variable itself is independent of the second value (property) or variable, conditional on that value of a third variable.

For example, consider the variables **SEX** [Male, Female] and the two properties Right-handed and Blond. If Male _||_ Right-handed | Blond, then Female _||_ Right-handed | Blond, which means that **SEX** _||_ Right-handed | Blond (since every value of **SEX** is independent of Right-handed conditional on Blond).

This fact about binary variables is especially useful when we are considering whether two binary variables are independent conditional on some property. Consider the variables: **SEX** [Male, Female], and **SMOKES** [Smoker, Non-smoker] and the property Student. In section 4320, we saw that Male _||_ Smoker | Student implies three other independencies. But this one independence among properties also implies the variable independence: **SEX** _||_ **SMOKES** | Student, since every value of **SEX** is independent of every value of **SMOKES** conditional on the property Student.

+ Male _||_ Smoker | Student if and only if
+ Male _||_ Non-Smoker | Student if and only if
+ Female _||_ Smoker | Student if and only if
+ Female _||_ Non-smoker

Thus, from Male _||_ Smoker | Student we know that **SEX** _||_ **SMOKES** | Student.

Suppose the variable **STATUS** can take on values [Student, Non-student]. Does Male _||_ Smoker | Student mean that **SEX** _||_ **SMOKES** | **STATUS**? No. Remember from section 5320 that Male _||_ Smoker | Student **does not mean that** Male _||_ Smoker | Non-student. Thus, just from knowing that Male _||_ Smoker | Student we do not know whether **SEX** _||_ **SMOKES** | **STATUS**.

In order to know whether the **variable** **SEX** is independent of the **variable** **SMOKES** conditional on the variable **STATUS**, we would have to check the independence of Male and Smoker (or two other properties) for **both** values of **STATUS**. If we did know both that Male _||_ Smoker | Student **and** Male _||_ Smoker | Non-student, then we could conclude that **SEX** _||_ **SMOKES** | **STATUS**.

---

## 6000: Summary

---

Two properties A and B are independent conditional on a third property C when, in the sub-population of individuals with C, learning about whether an individual has A does not change the chances that the individual has B.

In terms of relative frequencies, this means that A and B are independent conditional on C when the frequency of A in the sub-population of individuals with C is the same as the frequency of A in the sub-population of individuals with both B and C. Formally, A is independent of B conditional on C when Fr(A | C) = Fr(A | B, C).

$$\mathbf{A} \ \_||\__{S} \ \mathbf{B} \mid \mathbf{C}$$

if and only if

$$Fr_{S}(\mathbf{A} \mid \mathbf{C}) = Fr_{S}(\mathbf{A} \mid \mathbf{C}, \mathbf{B})$$

FIGURE 6000-1

Alternately, A _||_ B | C if and only if Fr(A & B | C) = Fr(A | C) * Fr(B | C). So, when we need to know whether two properties are independent conditional on a third property, we either check to see whether these two conditional frequencies are equal (by looking at histograms, or computing frequencies from a data table or contingency table), or else we check whether the conditional frequency of the complex property is just the product of the frequencies of the simple properties. Two properties are conditionally associated if they are not conditionally independent.

Like unconditional independence, conditional independence (and so also conditional association) is symmetric; if A is independent of (or associated with) B conditional on C, then B is independent of (or associated with) A conditional on C. Also, if two properties are independent conditional on a third property, then their complements are also independent (of the other property and the other complement) conditional on that third property. So, if we learn that A and B are independent (or associated) conditional on C, then we automatically learn three other independencies (associations).

It is harder to learn whether variables are independent conditional on some property, though, because every value of one variable must be independent of every value of the other variable conditional on that property. So, we typically have quite a few independencies that must be true. Values of binary variables, though, are just properties and their complements. So, learning that one value of a binary variable is independent of a property or another variable conditional on some property C implies that the binary variable itself is independent of the other property or variable conditional on that property C.

---