out to be [.3071, .4975]. The 97.5% CI for π_{sh} is [.1815, .3525] and the CI for π_{hb} is [.0853, .2231]. From Formula 4.8, the lower limit of the CI for $T(A)$ is .1815 and the upper limit is .3525. Whereas the lower limit differs substantially from 0, the upper limit suggests that this set is at most moderately fuzzy.

5. SIMPLE RELATIONS BETWEEN FUZZY SETS

5.1 Intersection, Union, and Inclusion

This chapter focuses on three of the elementary relationships offered by fuzzy set theory that are arguably distinctive but unfamiliar members of the family of bivariate associations (e.g., correlation, odds-ratio). These are fuzzy intersection, union, and inclusion. As explained in Chapter 2, the conventional rules for evaluating the membership of x in the intersection and union of fuzzy sets A and B are

$$\begin{aligned} m_{A\cap B}(x) &= \min(m_A(x), m_B(x)) \text{ and} \\ m_{A\cup B}(x) &= \max(m_A(x), m_B(x)). \end{aligned} \qquad [5.1]$$

Likewise, the rule stipulating that the fuzzy set A includes B ($A \supset B$) is that for all x,

$$m_A(x) \geq m_B(x). \qquad [5.2]$$

Intersection and union are distinct from addition because they are not compensatory, whereas addition is. For example, for any x from Formula 5.1, we can see that a high degree of membership $m_A(x)$ in A will not compensate a low membership $m_B(x)$ in B regarding membership $m_{A\cap B}(x)$ in $A\cap B$. Inclusion is distinct from correlation both because of its asymmetry (i.e., the extent to which A includes B tells us little about the extent to which B includes A) and its direct relationship with the logical concepts of necessity and sufficiency.

As mentioned in Chapter 3, one of the chief requirements for evaluating intersection, union, or inclusion empirically is *property ranking* (e.g., does Japan have higher membership in the set of "Asian countries" than in the set of "capitalist economies"?). Accordingly, we shall be concerned not only with the level of measurement but also with property ranking when considering techniques for evaluating fuzzy intersection or inclusion. For the remainder of this section, however, we will explore two illustrative examples in which property ranking may reasonably be assumed. Smithson (2005) goes over additional examples in detail.

➢ EXAMPLE 5.1: Attitudes Toward Immigrants

Fuzzy set inclusion is a generalization of crisp set inclusion and thereby conceptually related to Guttman, Mokken, and Rasch scaling. Although the inequality in Formula 5.2 seldom is perfectly satisfied, real examples may be found where it holds to quite a high degree. Figure 5.1 shows one such instance, in which 84 second-year psychology students at the Australian National University rated their degree of agreement with the propositions A = "Australia should permit immigrants to enter the country" and B = "Boat people should be allowed to enter Australia and have their claims processed." There are only three exceptions to the inclusion relationship $A \supset B$, in the upper right-hand corner of the scatterplot.

This example illustrates an important connection between intersection, union, and inclusion. If $A \supset B$, then $A \cap B$ is identical to the smallest set A or B, and $A \cup B$ is identical to the larger of the two sets. In Figure 5.1, we can see that the membership assignments for $A \cap B$ will be $m_{A \cap B}(x) = m_B(x)$ in all but the three cases where $m_B(x) = 1$ and $m_A(x) = 5/6$. Likewise, we can see that the membership assignments for $A \cup B$ will be $m_{A \cup B}(x) = m_A(x)$ in all but the same three cases. The nearer the distribution of membership values $m_{A \cap B}(x)$ to those of the smallest of A or B, the closer the relationship between A and B to a true inclusion relationship. The same is true regarding the distribution of $m_{A \cup B}(x)$ and those of the largest of A or B. Of course, computing intersections and unions depends on property ranking. Without good reason, it is unwise simply to assume that property ranking holds. We feel comfortable with it here because the response scale for each item is the same and the items are of the same form.

The example in Figure 5.1 also highlights a connection between fuzzy inclusion and the logical concepts of *necessity* and *sufficiency*. A predictive interpretation of the scatterplot in Figure 5.1 is that a high membership in Set A is necessary but not sufficient to predict high membership in Set B (or conversely, high membership in B is sufficient but not necessary to predict high membership in A). These asymmetric logical or predictive relations are not assessable by symmetric measures of association such as correlation. To characterize the pattern in Figure 5.1 by saying that the two

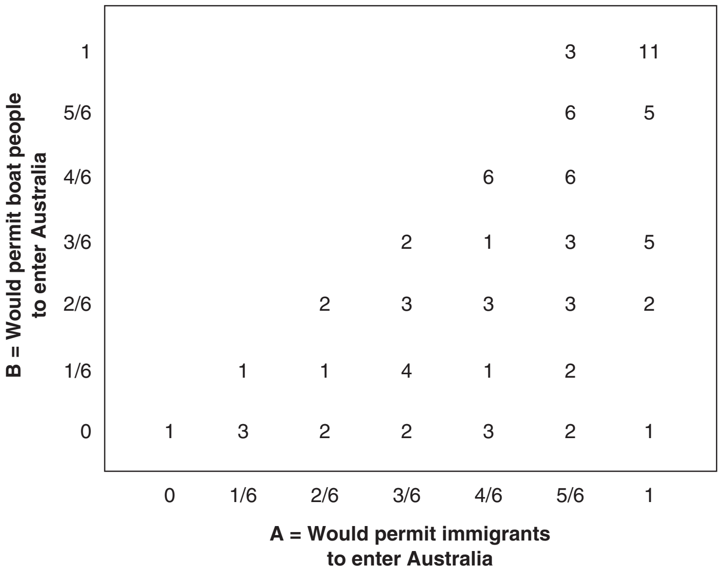

Figure 5.1 Example of Fuzzy Set Inclusion

variables have a correlation of .299 would surely miss the point. Even pointing out heteroscedasticity would not be specific enough.

Finally, it is noteworthy that inclusion, necessity, and sufficiency are special cases of a very useful and broad class of relations called *fuzzy restrictions*. A third interpretation of Figure 5.1 is that the joint distribution of A and B almost strictly satisfies the inequality $m_A(x) - m_B(x) \geq 0$. Fuzzy restrictions are generalizations of such inequalities.

➢ Example 5.2: Decisions to Disclose or Withhold Information

In an occupational survey (Bopping, 2003), 229 respondents were presented with a dilemma over whether to disclose information confided to them by a colleague. The respondents rated each of these statements on two identical 7-point scales:

I = "It is important to provide information to others"
T = "It is important to maintain the trust of a confidant"

TABLE 5.1
Cross-Tabulation of *I* and *T*

		m_I							
		0	*1/6*	*2/6*	*3/6*	*4/6*	*5/6*	*1*	*Total*
	1	9	4	5	8	4	6	23	59
	5/6		8	9	7	11	22	5	62
	4/6		2	14		4	15	1	36
m_T	*3/6*			1	7	6	10	9	33
	2/6		1	1		4	8	3	17
	1/6		1			1	5	4	11
	0	1				1		9	11
	Total	10	16	30	22	31	66	54	229

		I∪T							
		0	*1/6*	*2/6*	*3/6*	*4/6*	*5/6*	*1*	*Total*
observed	*pdf*	1	1	2	8	32	95	90	229
expected	*pdf*	0.48	2.02	7.04	15	26.9	78.5	99.1	229
sq. std.	*resid.*	0.56	0.51	3.61	3.26	0.97	3.47	0.83	13.2

We will demonstrate the application of fuzzy sets to investigating the hypothesis that respondents produced high ratings on *I* or *T* or both.

This hypothesis may be interpreted as saying that membership in the fuzzy union $I \cup T$ should be strongly skewed toward 1. A "stronger" version would predict that $T \supset \sim I$ (or equivalently, $I \supset \sim T$), i.e., $m_I(x) \geq 1 - m_T(x)$. An equivalent fuzzy restriction is $m_I(x) + m_T(x) \geq 1$. The scales for *I* and *T* have identical response formats, so for the sake of illustration, we will assume that the property ranking issue is resolved. The data in the upper part of Table 5.1 show that the strong version of this hypothesis is true for all but nine cases.

A comparison of the observed $I \cup T$ distribution with its expected-values counterpart if *I* and *T* are assumed independent (bottom part of Table 5.1) suggests that it is more strongly negatively skewed than would be expected under independence. A chi-square test may be used to compare the two distributions, and the squared standardized residuals constituting the chi-square statistic are shown in the third row of the bottom table. The chi-square test yields $\chi^2(6) = 13.2$, $p = .04$, thereby supporting the skew hypothesis. As in Example 5.1, the bivariate hypothesis investigated here would be rather difficult to evaluate using the usual concepts and measures of association but is readily accessible via fuzzy sets.

5.2 Detecting and Evaluating Fuzzy Inclusion

The task of detecting and evaluating fuzzy inclusion raises three questions. First, how do we assess the degree to which the fuzzy inclusion $m_A(x) \geq m_B(x)$ rule is satisfied? Second, how can we distinguish fuzzy inclusion from "impostors" such as the bivariate distribution of two independent skewed variables? And third, when do we have grounds for preferring a fuzzy set interpretation of our findings to rival interpretations?

Beginning with the first question, a number of fuzzy set theorists (e.g., Dubois & Prade, 1980, p. 22) have criticized the $m_A(x) \geq m_B(x)$ rule as too inflexible, and not sufficiently fuzzy. Smithson's (1987, pp. 31–32, 101–104) review of alternative proposals for evaluating fuzzy inclusion finds that they fall into two groups. One approach is to "fuzzify" the $m_A(x) \geq m_B(x)$ rule (e.g., Dubois & Prade, 1980; Ragin, 2000). The other is to construct an index of the degree of inclusion based on fuzzy set operators or other appropriate concepts. Both approaches hinge on the level of measurement possessed by the membership scales. We defer the discussion of this issue to the next section. Instead, we turn to the questions of distinguishing inclusion from impostors and deciding whether a bivariate relationship is better described as inclusion or some other kind of association. Table 5.2 makes this point by showing three impostors (the first, third, and fourth tables) and a genuine inclusion relationship (the second table).

➢ Example 5.3: Realistic Job-Seeking/Avoiding Example

The second table is taken from real data (Smithson & Hesketh, 1998), namely 360 respondents' responses to two items on the Holland vocational

TABLE 5.2
Inclusion Relation and Impostors

Independence + Skew

	Seek							
Not Avoid	*0*	$m_S(2)$	$m_S(3)$	$m_S(4)$	$m_S(5)$	$m_S(6)$	*1*	*Total*
0	4	3	1	1	1	0	0	10
$m_{NA}(2)$	5	4	2	2	1	1	0	15
$m_{NA}(3)$	6	4	3	2	1	1	0	17
$m_{NA}(4)$	7	5	3	2	1	1	1	20
$m_{NA}(5)$	20	13	8	6	4	2	2	55
$m_{NA}(6)$	33	23	12	10	6	4	3	91
1	55	38	21	17	11	6	4	152
Total	130	90	50	40	25	15	10	360

Inclusion Relationship

	Seek							
Not Avoid	*0*	$m_S(2)$	$m_S(3)$	$m_S(4)$	$m_S(5)$	$m_S(6)$	*1*	*Total*
0	8	3	2	1	1	1	0	16
$m_{NA}(2)$	8	4	3	2	0	0	0	17
$m_{NA}(3)$	17	11	7	10	1	0	0	46
$m_{NA}(4)$	13	11	22	30	6	0	1	83
$m_{NA}(5)$	5	7	12	23	3	2	0	52
$m_{NA}(6)$	1	5	10	25	19	9	1	70
1	3	3	2	16	13	13	26	76
Total	55	44	58	107	43	25	28	360

Positive Correlation

	Seek							
Not Avoid	*0*	$m_S(2)$	$m_S(3)$	$m_S(4)$	$m_S(5)$	$m_S(6)$	*1*	*Total*
0	47	0	0	0	0	0	0	47
$m_{NA}(2)$	0	47	5	0	0	0	0	52
$m_{NA}(3)$	0	0	47	10	4	0	0	61
$m_{NA}(4)$	0	0	0	45	10	0	0	55
$m_{NA}(5)$	0	0	0	0	46	5	0	51
$m_{NA}(6)$	0	0	0	0	0	47	0	47
1	0	0	0	0	0	0	47	47
Total	47	47	52	55	60	52	47	360

Negative Correlation

	Seek							
Not Avoid	*0*	$m_S(2)$	$m_S(3)$	$m_S(4)$	$m_S(5)$	$m_S(6)$	*1*	*Total*
0							16	16
$m_{NA}(2)$						4		4
$m_{NA}(3)$					14			14
$m_{NA}(4)$				87				87
$m_{NA}(5)$			92					92
$m_{NA}(6)$		92						92
1	55							55
Total	55	92	92	87	14	4	16	360

interest inventory. One item has them rate the extent to which they would *seek* a job that involves "realistic" tasks, and another asks them to rate the extent to which they would *avoid* this kind of job. Both scales were identical (ranging from *not at all* to *very strongly*), and the "avoid" scale has been reverse-scored to convert it into a "not avoid" scale. The hypothesized relationship is that seeking this kind of job is sufficient but not necessary to also not avoid it, because one could decide not to avoid it for other reasons as well. So "seeking" is included in "not avoiding."

All four two-way tables have a very similar proportion of cases obeying the fuzzy inclusion rule $m_A(x) \geq m_B(x)$. Excluding the zero-membership cases on the included set, the proportions are .887, .889, .891, and .889 for the first, second, third, and fourth tables, respectively. However, the uppermost table was generated by cross-tabulating two independent skewed distributions. A chi-square test for this table yields $\chi^2(36) = 3.669$, which is a very good fit with the independence model. The apparently strong inclusion relationship in this table is due solely to the skew in both distributions.

Moving now to the other three tables in Table 5.2, a chi-square test yields $\chi^2(36) = 234.036$ for the second table, $\chi^2(36) = 1781.344$ for the third table, and $\chi^2(36) = 3625.220$ for the fourth table, all indicating large departures from independence. However, the third and fourth tables show very strong correlation patterns rather than an inclusion relationship, even though the proportion of cases obeying the fuzzy inclusion rule is nearly identical to the second table. Many researchers would prefer to describe the third and fourth tables in terms of this correlation, which measures the strength of a one-to-one association between two variables, as opposed to the one-to-many of necessity. We could readily imagine (and find) "intermediate" situations in which there is both a moderately strong correlation and a reasonably strong inclusion relationship.

Which description should we prefer, and why? This problem is more difficult than merely detecting independence, because more judgments are required. For instance, even if correlation provides a "good" description of the relationship (i.e., all assumptions and requirements such as homoscedasticity are satisfied), the inclusion interpretation might still be the more theoretically relevant. On the other hand, inclusion is a one-to-many relation and thus is a less precise proposition than a one-to-one relationship such as correlation or stronger measures of association.

Let us dispense with independence + skew first. Independence + skew cannot be a genuine inclusion relation because there *is* no association between two statistically independent random variables. Nevertheless, it is easy to make examples of skewed statistically independent random variables that seem to satisfy fuzzy inclusion, as our example demonstrates.

When two variables are statistically independent, their joint distribution is *completely* determined by the marginal distributions because the joint distribution is just the product of the marginals. The marginals, in turn, depend on the assignment of membership. As we have already seen in Chapter 3, assignment of membership is a very difficult task. It is wise to rest one's conclusions on the assignment process as little as possible because it is almost always possible to argue that a given assignment is wrong. For discrete membership scales, as in this example, the conventional chi-square test of independence usually will suffice. For continuous membership scales, the Kolmogorov-Smirnov test is the most well-known, and it compares the observed joint cumulative distribution function (JCDF) against the expected JCDF under independence. For alternative approaches, see D'Agostino and Stephens (1986).

Association + skew raises other issues. We take the view that if the bivariate distribution satisfies the relevant assumptions and the researcher is primarily interested in predicting one variable from the other, then a correlation-regression description may be preferable to a fuzzy set perspective. Better still would be a GLM that models location and dispersion simultaneously. On the other hand, particular kinds of heteroscedasticity, a strong inclusion rate combined with a marked difference in the sizes of the two sets, and/or research questions that are expressed in set-like terms should motivate a serious consideration of fuzzy inclusion as a description of the patterns. The next two sections present techniques for investigating inclusion relations in detail.

5.3 Quantifying and Modeling Inclusion: Ordinal Membership Scales

In many circumstances, we may wish to evaluate how robust a claim about an inclusion relationship is against alternative membership value assignments. For both the $m_A(x) \geq m_B(x)$ rule and any inclusion index, the joint ordering of membership values for the two sets crucially determines the result, so it is essential to explore what happens to inclusion rates and index values when the joint ordering is modified. A reasonable approach to assessing how dependent our results are on the joint ordering of membership values is to stipulate a benchmark inclusion rate before seeing the data, and then ascertain the collection of paths whose confidence intervals (CIs) include that rate or higher. One way to determine the relevant "collection" is to begin with a specific joint ordering of the values that yields a path whose inclusion CI contains the prescribed rate, and then ascertain which neighboring paths' CIs also include that rate.

TABLE 5.3
Paths With Inclusion Rate CI Containing 0.9

	Seek							
Not Avoid	*0*	m_S*(2)*	m_S*(3)*	m_S*(4)*	m_S*(5)*	m_S*(6)*	*1*	*Total*
0	8	3	2	1	1	1	0	16
m_{NA}*(2)*	8	4	3	2	0	0	0	17
m_{NA}*(3)*	17	11	7	10	1	0	0	46
m_{NA}*(4)*	13	11	22	30	6	0	1	83
m_{NA}*(5)*	5	7	12	23	3	2	0	52
m_{NA}*(6)*	1	5	10	25	19	9	1	70
1	3	3	2	16	13	13	26	76
Total	55	44	58	107	43	25	28	360

To see how this works, let us return to the job-seeking example using a criterion inclusion rate of .9. As mentioned earlier, the proportion of cases obeying the $m_A(x) \geq m_B(x)$ rule for the diagonal path is 271/305 = .889. A 95% CI for this path is [.848, .922], so it is compatible with an inclusion rate of .9. In fact, it can be shown that any path with a proportion as low as 264/305 has a CI that includes .9.

The second table from Table 5.2 is reproduced in Table 5.3. The shaded region denotes the collection of paths involving one alteration in the original values' joint ordering whose inclusion proportions are at least 264/305. The joint ordering of memberships corresponding to the diagonal path is $0 < m_S(2) = m_{NA}(2) < m_S(3) = m_{NA}(3) < m_S(4) = m_{NA}(4) < m_S(5) = m_{NA}(5) < m_S(6) = m_{NA}(6) < 1$. The collection of paths forms a region that begins only slightly beneath the diagonal path. For example, the path deviating from the diagonal once by following the cells with frequencies {8, 4, 22, 30, 3, 9, 26} corresponds to the joint ordering $0 < m_S(2) = m_{NA}(2) < m_{NA}(3) < m_S(3) < m_S(4) = m_{NA}(4) < m_S(5) = m_{NA}(5) < m_S(6) = m_{NA}(6) < 1$.

In the absence of an inclusion rate criterion, we may use the single alteration in the joint-ordering criterion for exploring the sensitivity of inclusion rates to membership assignments. Again starting with the inclusion rate of .889 for the diagonal path, the largest possible change in this rate incurred by one alteration of the joint ordering is the exclusion of the 30 cases in the $\{m_S(4) = m_{NA}(4)\}$ cell. Excluding them by "lowering" the path decreases the inclusion rate from .889 to (271 – 30)/305 = .790. The biggest increase of the inclusion rate by one alteration in joint ordering is the inclusion of the 10 cases in the $\{m_S(4) = m_{NA}(3)\}$ cell, resulting in a rate of (217 + 10)/305 = .921. Neither the proportion of cases obeying the $m_A(x) \geq m_B(x)$

rule nor the inclusion indexes do a good job of distinguishing between negative correlation and genuine inclusion because they are strongly influenced by the marginal distributions. The tests proposed in Ragin (2000) based on the $m_A(x) \geq m_B(x)$ rule implicitly assume that the marginal distributions are uniform. To avoid making strong assumptions about the marginal distributions, we must turn to models of inclusion based on localized inclusion relations in tables and scatterplots.

One way of modeling inclusion throughout a scatterplot or table is via level sets, which were introduced in Chapter 2. We may establish the inclusion rate for any cell in a table (or point on a scatterplot) by constructing the joint cumulative distribution function (JCDF). The first table in Table 5.4 shows the JCDF from Table 5.3, which accumulates frequencies starting in the {1,1} cell at the lower right and moving upward and to the left. That cell contains 26 cases, so moving up one cell accumulates 1 more to give 26 + 1 = 27, whereas moving one cell to the left accumulates 13 cases to give 26 + 13 = 39, moving one cell up and to the left accumulates 1 + 13 + 9 cases to give 26 + 1 + 13 + 9 = 49, and so on.

The second table shows the local inclusion rate for each cell. These are determined by dividing the cumulative frequency in that cell by the column cumulative total, located in the first row of the table. For the lower-right cell, we have 26/28 = .929, for the next cell to the left we have 39/53 = .736, and so on. We may regard these proportions as local inclusion rates because each of them is the proportion of cases that obeys the $m_A(x) \geq m_B(x)$ rule for the level set that corresponds to its cell. Consider the {.83, .83} cell, which has a JCDF of 49 cases. There are 53 cases for which membership in Seek is .83 or above, and 49 of those obey the $m_A(x) \geq m_B(x)$ rule because their membership in Not Avoid also is .83 or above. The proportion is therefore 49/53 = .925, the entry for that cell in the second table.

The level set and JCDF approach enables researchers to examine patterns of local inclusion rates. Notice that the inclusion rates for the cells on the diagonal path are very similar to one another. This path arguably has a constant inclusion rate along it, and we shall see shortly how to test a constant inclusion model for this path. The inclusion rate pattern in Table 5.4 contrasts vividly with that for the negative correlation example, shown in Table 5.5. The inclusion rates along the diagonal path clearly are not constant, but instead jump suddenly from 0 to quite high levels as we move upward and to the left along that path. This comparison demonstrates that local inclusion models can distinguish between relationships that a global inclusion rate or inclusion index cannot.

Now let us test a constant inclusion model for the diagonal path in the job-seeking example. It turns out that the average inclusion rate along this path is .947. We will test whether the local inclusion pattern along the

TABLE 5.4
JCDF and Local Inclusion Rates for Example 5.3

Joint Cumulative Distribution

	Seek						
Not Avoid	*0*	m_S*(2)*	m_S*(3)*	m_S*(4)*	m_S*(5)*	m_S*(6)*	*1*
0	360	305	261	203	96	53	28
m_{NA}*(2)*	344	297	256	200	94	52	28
m_{NA}*(3)*	327	288	251	198	94	52	28
m_{NA}*(4)*	281	259	233	187	93	52	28
m_{NA}*(5)*	198	189	174	150	86	51	27
m_{NA}*(6)*	146	142	134	122	81	49	27
1	76	73	70	68	52	39	26

Local Inclusion Rates

	Seek						
Not Avoid	*0*	m_S*(2)*	m_S*(3)*	m_S*(4)*	m_S*(5)*	m_S*(6)*	*1*
0							
m_{NA}*(2)*	0.956	0.974	0.981	0.985	0.979	0.981	1.000
m_{NA}*(3)*	0.908	0.944	0.962	0.975	0.979	0.981	1.000
m_{NA}*(4)*	0.781	0.849	0.893	0.921	0.969	0.981	1.000
m_{NA}*(5)*	0.550	0.620	0.667	0.739	0.896	0.962	0.964
m_{NA}*(6)*	0.406	0.466	0.513	0.601	0.844	0.925	0.964
1	0.211	0.239	0.268	0.335	0.542	0.736	0.929

diagonal path is consistent with a constant inclusion rate of .947. There are several methods for doing this, but the most familiar and perhaps simplest is to use a chi-square test. The principle is to generate expected frequencies for the JCDF along the diagonal path, obtain expected frequencies for that path by taking the differences between adjacent cells, and then compare those with the observed frequencies for the same path using a one-way chi-square test.

The first table in Table 5.6 shows how to obtain the observed frequencies for each cell in the diagonal path by taking differences between adjacent cells, starting at the lower right. The second table shows how the expected frequencies are computed, using an inclusion rate of .947 and the marginal observed frequencies in the first row of the first table. The expected frequency for the upper-leftmost cell, 71.165, is computed by subtracting the sum of the other expected frequencies from the total sample size, 360.

TABLE 5.5
JCDF and Local Inclusion Rates for Negative Correlation Example

Joint Cumulative Distribution

	Seek						
Not Avoid	*0*	m_S*(2)*	m_S*(3)*	m_S*(4)*	m_S*(5)*	m_S*(6)*	*1*
0	360	305	213	121	34	20	16
m_{NA}*(2)*	344	289	197	105	18	4	0
m_{NA}*(3)*	340	285	193	101	14	0	0
m_{NA}*(4)*	326	271	179	87	0	0	0
m_{NA}*(5)*	239	184	92	0	0	0	0
m_{NA}*(6)*	147	92	0	0	0	0	0
1	55	0	0	0	0	0	0

Local Inclusion Rates

	Seek						
Not Avoid	*0*	m_S*(2)*	m_S*(3)*	m_S*(4)*	m_S*(5)*	m_S*(6)*	*1*
0							
m_{NA}*(2)*	0.956	0.948	0.925	0.868	0.529	0.200	0.000
m_{NA}*(3)*	0.944	0.934	0.906	0.835	0.412	0.000	0.000
m_{NA}*(4)*	0.906	0.889	0.840	0.719	0.000	0.000	0.000
m_{NA}*(5)*	0.664	0.603	0.432	0.000	0.000	0.000	0.000
m_{NA}*(6)*	0.408	0.302	0.000	0.000	0.000	0.000	0.000
1	0.153	0.000	0.000	0.000	0.000	0.000	0.000

Because there are seven cells, we have 6 degrees of freedom and so, using a significance criterion of .05, the critical chi-square value is 12.592. The observed chi-square turns out to be $\chi^2(6) = 3.257$, which is well below the critical value and indicates a rather good fit between the constant inclusion model and the data. There are inclusion rates other than .947 that we could not reject using the chi-square test. It is not difficult to obtain a 95% confidence interval for the constant inclusion rates, although it should be borne in mind that the chi-square version is conservative, and the resulting CI is [.888, 1]. Likewise, it is possible to find the collection of all paths that are compatible with a constant inclusion model, whether for a prespecified rate or in general. However, a full exploration of this topic is beyond the scope of this chapter.

Now let us test a constant inclusion model for the diagonal path in the negative correlation example. Table 5.7 shows the observed frequencies. It turns out that no matter what inclusion rate is used, the chi-square test

TABLE 5.6
Constant Inclusion Model

Observed Frequencies

	Seek						
Not Avoid	*0*	m_S*(2)*	m_S*(3)*	m_S*(4)*	m_S*(5)*	m_S*(6)*	*1*
0	360 − 305 = 55	44	58	107	43	25	28
m_{NA}*(2)*		297 − 251 = 46					
m_{NA}*(3)*			251 − 187 = 64				
m_{NA}*(4)*				187 − 86 = 101			
m_{NA}*(5)*					86 − 49 = 37		
m_{NA}*(6)*						49 − 26 = 23	
1							26

Expected Frequencies

	Seek						
Not Avoid	*0*	m_S*(2)*	m_S*(3)*	m_S*(4)*	m_S*(5)*	m_S*(6)*	*1*
0	71.165	44	58	107	43	25	28
m_{NA}*(2)*		.947 × 44 = 41.668					
m_{NA}*(3)*			.947 × 58 = 54.926				
m_{NA}*(4)*				.947 × 107 = 101.329			
m_{NA}*(5)*					.947 × 43 = 40.721		
m_{NA}*(6)*						.947 × 25 = 23.675	
1							.947 × 28 = 26.516

rejects a constant inclusion model for this table. The lowest chi-square obtainable (for an inclusion rate of 1) is $\chi^2(6) = 40.959$, substantially higher than the critical chi-square value of 12.592. The constant inclusion model successfully distinguishes between the job-seeking and negative correlation examples.

TABLE 5.7
Frequencies for the Negative Correlation Example

Observed Frequencies

	Seek						
Not Avoid	*0*	$m_S(2)$	$m_S(3)$	$m_S(4)$	$m_S(5)$	$m_S(6)$	*1*
0	360 – 289 = 71	92	92	87	14	4	16
$m_{NA}(2)$		289 – 193 = 96					
$m_{NA}(3)$			193 – 87 = 106				
$m_{NA}(4)$				87			
$m_{NA}(5)$					0		
$m_{NA}(6)$						0	
1							0

Likewise, it can be readily proved that the only constant inclusion paths for two statistically independent fuzzy sets are horizontal. For ordinal categorical membership functions and contingency tables, this property follows from the same argument behind the formula for computing expected frequencies when independence is assumed. The horizontal inclusion path result underscores our reasons for not considering the independence + skew inclusion pattern as a genuine inclusion relation.

5.4 Quantified and Comparable Membership Scales

When $m_A(x)$ and $m_B(x)$ are quantified and comparable, the fuzzy set tool chest opens up. This concluding section to Chapter 5 presents a brief survey of the possibilities that await the researcher in this situation.

➢ Example 5.4: Fear and Loathing in the Tropics

The data set that will be used for illustrations here was collected from 262 psychology undergraduates (from James Cook University in a tropical region in Australia). The data comprise their self-reported feelings about 31 noxious stimuli, such as snakes or vomit. They were asked to rate their degree of fear, disgust, and dislike for each stimulus, using a 4-point rating scale ranging from 0 = *not at all* to 3 = *very much.* The 31 ratings of fear, disgust, and dislike were summed and divided by 31 to obtain fuzzy membership scales for each. For our purposes here, we will treat these as quantified comparable membership scales. The chief object in this study was a hypothesis that the "phobic"-style responses fear and disgust are subsets of dislike, which is considered to be a much broader emotional response. A subsidiary question was the "comorbidity" issue raised often in clinical and health psychology; that is, to what extent respondents simultaneously fear and are disgusted by noxious stimuli.

5.4.1 Cardinality of Intersections and Unions

Starting with the comorbidity issue, a traditional approach would use correlation (we shall explore this issue in greater detail in Chapter 6). Table 5.8 shows that all three fuzzy sets are significantly and moderately correlated. However, the correlations cannot tell us whether one set strongly includes another, nor do they provide meaningful estimates of the relative sizes of these sets or their intersections.

We may measure the cardinality (size) of fuzzy set intersections and unions, thereby enabling us to directly address comorbidity. The upper subtable in Table 5.9 shows the mean memberships of *Fear*, *Dislike*, and *Disgust* on the diagonal and the mean memberships of their pairwise intersections in the off-diagonal cells. The lower subtable shows the proportion of each set accounted for by its intersection with another set. For instance, the intersection between *Fear* and *Dislike* has average membership .229. Because the average membership in *Fear* is .231 and in *Dislike* is .563, the proportion of *Fear* accounted for in the intersection is .229/.231 = .991 and the proportion of *Dislike* accounted for is .229/.563 = .407.

The comorbidity picture presented by fuzzy intersections differs vividly from the correlational perspective. It is evident that the comorbidity rate for *Fear* and *Disgust* is quite high (78.4% of *Fear* and 84.2% of *Disgust* are accounted for by their intersection). *Dislike* clearly includes most of *Fear* and *Disgust* (99.1% and 99.5% respectively), but only 38.0% of *Dislike* is accounted for by its intersection with *Disgust* and only 40.7% is accounted

TABLE 5.8
Correlations Among *Fear*, *Disgust*, and *Dislike*

Fear		
.434	*Dislike*	
.747	.410	*Disgust*

TABLE 5.9
Mean Membership of *Fear*, *Disgust*, *Dislike*, and Their Intersections

	Mean Membership		
	Fear	*Dislike*	*Disgust*
Fear	0.231		
Dislike	0.229	0.563	
Disgust	0.181	0.214	0.215
	Intersection Proportions		
	Fear	0.407	0.842
	0.991	*Dislike*	0.995
	0.784	0.380	*Disgust*

for by its intersection with *Fear*. The finding that *Dislike* strongly includes *Fear* and *Disgust* is supported by their scatterplots in Figure 5.2.

Even though *Dislike* subsumes most of *Fear* and *Disgust*, is the union of *Fear* and *Disgust* sufficient to include most of *Dislike*? The average membership in *Fear*∪*Disgust* turns out to be .266, which is less than half the size of *Dislike* (.563). In fact, the average membership in the intersection (*Fear*∪*Disgust*)∩*Dislike* is .263, so *Dislike* includes 100(.263/.266) = 98.7% of *Fear*∪*Disgust*. These findings indicate that *Dislike* is a much broader category than the union of *Fear* and *Disgust*. Even in this quick exploration of intersections and unions, we have gone far beyond anything that correlation or regression could tell us.

5.4.2 Inclusion Coefficients

We now turn to inclusion coefficients as a way of quantifying the degree to which one set includes another. The simplest index of inclusion is just the proportion of cases satisfying the $m_A(x) \geq m_B(x)$ rule, the "Classical Inclusion Ratio" presented in Chapter 2. While attractive for its simplicity, its main limitation is that a "near miss" is counted as strongly as a drastic counterexample. We present two coefficients that overcome this limitation, namely, the "Inclusion 1" and "Inclusion 5" indexes discussed in Smithson (1994). The first inclusion index is defined by

$$I_{AB} = \sum m_{A \cup B}(x_i) \Big/ \sum m_B(x_i). \qquad [5.3]$$

I_{AB} is the proportion of Set *B* in the intersection of Sets *A* and *B* (Sanchez, 1979). It is clearly based on fuzzy set-theoretic concepts. We already used

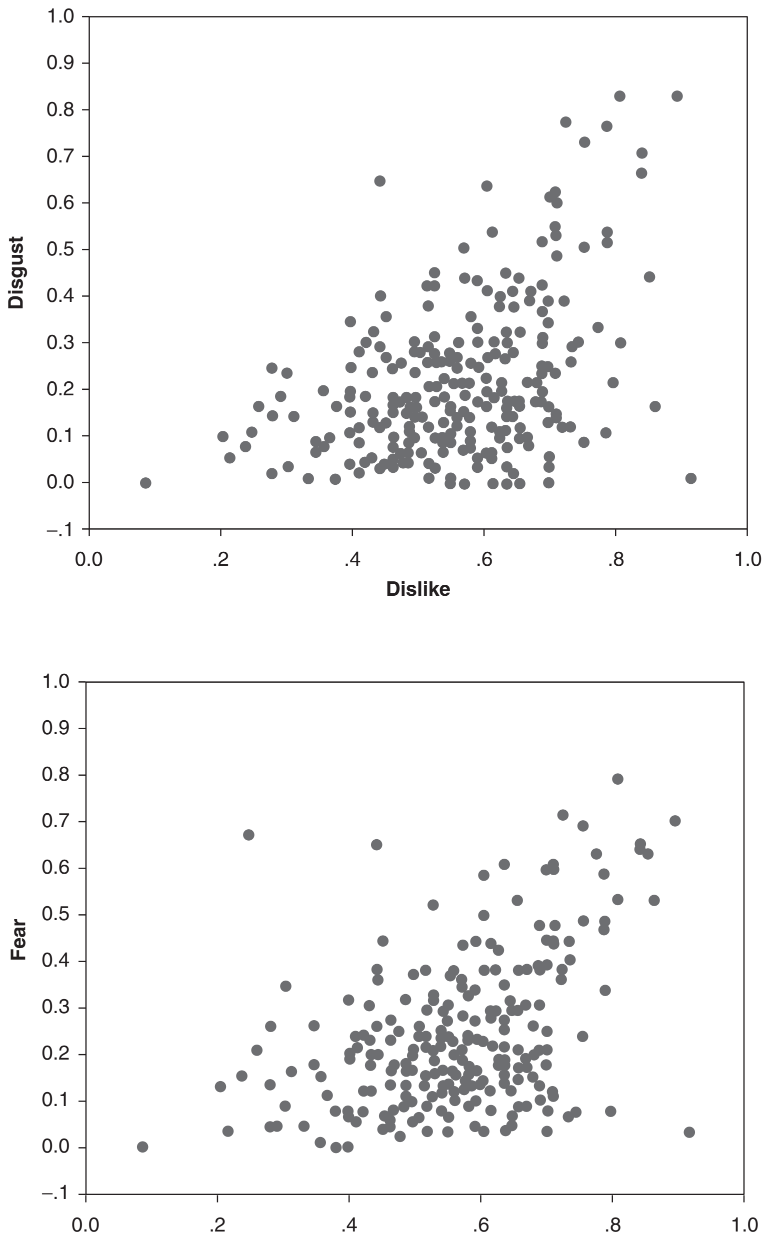

Figure 5.2 Scatterplots of *Disgust* × *Dislike* and *Fear* × *Dislike*

this coefficient in Table 5.9, where we discussed the proportion of one set accounted for by its intersection with another.

There is an important link between the JCDF and local inclusion approach from the previous section and *IAB*. Summing the JCDF values along the diagonal path in Table 5.4 and dividing by the number of cells gives (26 + 49 + 86 + 187 + 251 + 297)/6 = 149.333. This is the cardinality of the intersection between the *Seek* and *Not Avoid* sets, if we are willing to regard the membership scale as taking values of k/K, for $k = 0, 1, 2, \ldots, K$, where K is the number of nonzero membership levels. Summing the column CDF and dividing that by 6 yields (28 + 53 + 96 + 203 + 261 + 305)/6 = 157.667, the size of the *Seek* set. So the inclusion index $I_{AB} = 149.333/157.667 = .947$, and it may be thought of as the sum of the JCDF along the path divided by the sum of the CDF of the included set if we are willing to regard the membership values as taking values of k/K, where K is the number of cells in the path. This argument was the basis for choosing to test a constant inclusion model with a rate of .947.

The "Inclusion 5" coefficient is defined by

$$C_{AB} = \frac{\sum \max(0, m_A(x_i) - m_B(x_i))}{\sum |m_A(x_i) - m_B(x_i)|}. \qquad [5.4]$$

C_{AB} is the proportion of deviations from equality between $m_A(x)$ and $m_B(x)$ that are in the appropriate direction. It is actually a generalization of the proportion of observations with unequal membership obeying the strict inequality $m_A(x) > m_B(x)$.

Which of these indexes is preferable depends on the researcher's goals. To start with, cases of 0-valued membership for either set do not affect the value of I_{AB}, but they do affect C_{AB}. Second, cases where $m_A(x_i) = m_B(x_i)$ do not affect C_{AB}, but do affect I_{AB}. Third, $C_{AB} = 1 - C_{BA}$ but this does not hold for I_{AB}. Finally, neither coefficient is defined casewise, an attractive property for estimation purposes (Smithson, 1987, 1994 review others that are).

As with any coefficient designed to measure a particular kind of relationship and no other, inclusion coefficients have their limitations. First, neither of the inclusion coefficients tells us whether independence holds. Additionally, as mentioned previously, they are strongly influenced by the marginal distributions; an inclusion index that is free of the margins in the same way that the odds ratio is for 2 × 2 tables would be highly useful. Table 5.10 shows inclusion coefficients for the four tables in Table 5.2. For the independence example (the first table in Table 5.2), $I_{AB} = .914$ and $C_{AB} = .962$, both of which appear impressive unless we know about independence.