Journal of Probability and Statistical Science
21(2), 62-72 Sep. 2023

The Coefficient of Dependence and Conditioning

Stephen M. Scariano
Sam Houston State University

Ananda B. W. Manage
Sam Houston State University

ABSTRACT

The Coefficient of Dependence is introduced as an avenue for aiding student understanding of dependent, statistical events. The essential features of this coefficient are studied using different models and explored with multiple examples. Conditional probabilities are ultimately understood to be simple transformations of marginal probabilities via the Coefficient of Dependence, which is an idea well within the grasp of all undergraduate students. For completeness, proofs for the main results are relegated to the appendix for those interested.

Keywords: independence, conditioning, dependence, marginal, joint

Introduction

Student understanding of the whys and wherefores of conditional probability are intrinsically grounded in the interplay between the topics of independence and dependence. Our experience is that unless the latter two concepts are fully grasped, there is little hope that students truly perceive, and can use, the vast power of conditioning in statistical problem-solving.

Revisiting foundational concepts in the study of probability and statistics can occasionally inspire greater insights for teaching and learning. Initially, we open fresh eyes to the concept of independence with the goal of exploring new avenues for teaching this essential topic. Then, we delve into some surprising linkages with the general concept of conditioning.

1. The Coefficient of Dependence

First, consider the standard definition of independence found in many common texts.

Standard Definition: Let ***S*** be the sample space of a random experiment having probability function P defined for all events associated with ***S.*** Events A and B are said to be (probabilistically) ***independent*** when and only when

$$P(A \cap B) = P(A) \times P(B). \tag{1}$$

Note that $P(A) = 0$ or $P(B) = 0$ implies $P(A \cap B) = 0,$ inasmuch as $A \cap B$ is a subset of both events A and $B.$ Moreover, if $P(A \cap B) = 0,$ then equation (1) prevails if and only if $P(A) = 0$ or $P(B) = 0,$ These are the simplest cases for independence.

Returning to equation (1) as an exemplar, there are three pertinent probability statements regarding the relationship among the probabilities shown that can, and do, emerge in practice.

- Received September 2023, in final form September 2023.
- Stephen M. Scariano (corresponding author), Sam Houston State University, Huntsville, Texas 77340 Taxus,. scariano@shsu.edu and Ananda B. W. Manage, Sam Houston State University, Huntsville, Texas 77340 Taxus,. wxb001@shsu.edu

These are: $P(A \cap B) > P(A) \times P(B)$ or $P(A \cap B) = P(A) \times P(B)$, or $P(A \cap B) < P(A) \times P(B)$. Let's study these three cases more closely with an eye toward enhancing teaching effectiveness.

Example 1.1: Starting simply, consider the following three probability models associated with an unbalanced (biased) tetrahedra die to be tossed once with its downward face recorded,

Model I				
Face	1	2	3	4
Probability	0.4	0.1	0.1	0.4

Model II:				
Face	1	2	3	4
Probability	0.1	0.4	0.4	0.1

Model III:				
Face	1	2	3	4
Probability	0.16	0.24	0.24	0.36

Consider events $A = \{1,2\}$ and $B = \{1,3\}$, so that $A \cap B = \{1\}$.

For **Model I**: $P(A) = 0.50,\ P(B) = 0.50,$ and $P(A \cap B) = 0.40,$ so that

$$P(A \cap B) = \mathbf{0.40} > 0.25 = \mathbf{0.50} \times \mathbf{0.50} = P(A) \times P(B),$$

showing that events A and B are not independent (= dependent).

For **Model II**: $P(A) = 0.50,\ P(B) = 0.50,$ and $P(A \cap B) = 0.10,$ so that

$$P(A \cap B) = \mathbf{0.10} < 0.25 = \mathbf{0.50} \times \mathbf{0.50} = P(A) \times P(B),$$

showing that events A and B are not independent, or dependent.

And, for **Model III**: $P(A) = 0.40,\ P(B) = 0.40,$ and $P(A \cap B) = 0.16,$

$$P(A \cap B) = \mathbf{0.16} = \mathbf{0.40} \times \mathbf{0.40} = P(A) \times P(B),$$

showing that events A and B are independent. ■

Notice that for **Model III**, events A and B have a non-null intersection, $A \cap B = \{1\}$, and yet those two events are still independent, exemplifying the well-known fact that non-null independent events must overlap. So, the independence property does not imply disjointedness, which is a common student misconception. More importantly, **Models I** and **II** reveal different types (> vs. <) of dependence. This motivates the following definition which is the cornerstone of all that follows.

Definition 1.1: Let ***S*** be the sample space of a random experiment having probability function P defined for all events associated with ***S.*** Let A and B be any two non-null $(P(A) \neq 0 \text{ and } P(B) \neq 0)$ events. The **coefficient of dependence**, written $\mathbf{D}(A,B)$, is defined by

$$\mathbf{D}(A,B) = \frac{P(A \cap B)}{P(A) \times P(B)}. \tag{2}$$

Events A and B are said to be **independent** when and only when $\mathbf{D}(A,B) = 1$. Otherwise, A and B are said to be **super-dependent** when $\mathbf{D}(A,B) > 1$ or **sub-dependent** when $\mathbf{D}(A,B) < 1$.

Observe that the coefficient of dependence is a symmetric function of its two arguments in the sense

$$\mathbf{D}(B,A) = \frac{P(B \cap A)}{P(B) \times P(A)} = \frac{P(A \cap B)}{P(A) \times P(B)} = \mathbf{D}(A,B), \tag{3}$$

Additionally, $\mathbf{D}(A,B)$ is the ratio of $P(A \cap B)$ to $[P(A) \times P(B)]$, which is immediately seen to be a nonnegative, real number, whenever A and B are both non-null events.
The reason for studying the coefficient of dependence will be seen in the next section, which considers dependent events and their relationship to conditional probability statements. For now, simply note that events A and B are super-dependent ($\mathbf{D}(A,B) = 1.60$) for **Model I**, while sub-dependent ($\mathbf{D}(A,B) = 0.40$) for **Model II**, and yet independent ($\mathbf{D}(A,B) = 1.00$) for **Model III**.

2. Connections with Conditional Probability

The topic of conditional probability is a fundamental concept in the study of probability and statistics, and many modern discussions of it begin with the following definition.

Definition 2.1: Let ***S*** be the sample space of a random experiment having probability function P defined for all events associated with ***S.*** Let A and B be any two events. The conditional probability of event A given that event B has already occurred, written $P(A \mid B)$, is defined by

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \tag{4}$$

provided B is a non-null event.

A close connection between the coefficient of dependence and conditional probability can now be established. That is,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \left(\frac{P(A)}{P(A)}\right)\left[\frac{P(A \cap B)}{P(B)}\right] = \left[\frac{P(A \cap B)}{P(A)P(B)}\right] \times P(A) = \mathbf{D}(A,B) \times P(A) \tag{5a}$$

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \left[\frac{P(B \cap A)}{P(A)}\right]\left(\frac{P(B)}{P(B)}\right) = \left[\frac{P(A \cap B)}{P(A)P(B)}\right] \times P(B) = \mathbf{D}(A,B) \times P(B), \tag{5b}$$

when A and B are non-null events. Equations (5a) and (5b) reveal **precisely how** the marginal (unconditional) probabilities $P(A)$ and $P(B)$, are **transformed** into the conditional probabilities $P(A \mid B)$ and $P(B \mid A)$ when additional information is provided. In fact, the transformation is exclusively **quantified** by the coefficient of dependence factor $\mathbf{D}(A,B)$ **in both cases**. That is,

Transformation

$$P(A \mid B) = [\mathbf{D}(A,B)] \times P(A)$$
$$P(B \mid A) = [\mathbf{D}(A,B)] \times P(B)$$

Transformation

(6)

So, the conditional probabilities on the left-hand sides of equations (5) are both *transformed* (magnified or diminished, or unaltered) from their corresponding marginal probabilities by the same factor, the coefficient of dependence.

It is critically important here to keep in mind that equations (6) are ***not*** intended to be computationally simpler than defining equation (4), for they are not. Rather, equations (6) give the necessary *insight* into the transformative nature of conditional probability in a functional format quite familiar to students. Indeed, the acquisition of deeper student insight far overshadows any loss of computational simplicity, as the next example illustrates.

Example 2.1: (Written in the format of a classroom exercise)
The Information Technology (IT) department for a large, multinational corporation maintains two classifications, A and B, for its one hundred systems engineers. There are thirty A-classified systems engineers, and the remainder are B-classified. Sixty percent of the A-classified systems engineers are women, while twenty percent of the B-classified systems engineers are women.

(a) Write a cross-classification table using the given information.

	***A*-classified**	***B*-classified**	Row Total
Woman	18	14	32
Man	12	56	68
Column Total	30	70	100
			Grand Total

Consider the following events for selecting a systems engineer:

A = event that a selected person is A-classified. B = event that a selected person is B-classified.
M = event that a selected person is a man. W = event that a selected person is a woman.

(b) Suppose a systems engineer from the IT department of the company is randomly selected, find the probability that person is a woman.

$$P(W) = 0.32,\ \text{computed directly using the first-row total in the table.}$$

(c) Suppose a system engineer from the IT department of the company is randomly selected, find the probability that person is A-classified.

$$P(A) = 0.30,\ \text{computed directly using the first-column total in the table.}$$

(d) Suppose a system engineer from the IT department of the company is randomly selected, find the probability that person is an A-classified women.

$$P(W \cap A) = 0.18,\ \text{computed using the first-row, first-column cell of the table.}$$

(e) Find the coefficient of dependence factor, $\mathbf{D}(W, A)$.

$$\text{Using (b), (c) and (d) above, } \mathbf{D}(W, A) = \frac{P(W \cap A)}{P(W) \times P(A)} = \frac{0.18}{(0.32 \times 0.30)} = \mathbf{1.875}.$$

(f) Find the conditional probability a randomly selected person is a woman, given they are A-classified.

$$P(W \mid A) = \mathbf{D}(W, A) \times P(W) = \mathbf{1.875} \times 32\% = 60\%,\ \text{computed using equations (5).}$$

Events W and A are super-dependent since $P(W)$ = 32% almost doubles to $P(W| A)$ = 60% with the given information. Indeed, it is the coefficient of dependence factor, **1.875**, that provides the precise, transformational effect that the conditional information conveys beyond the marginal probability to the conditional probability. That is, the coefficient of dependence "mathematizes", using a single number, the given information to produce a conditional probability carrying that information from the marginal probability. In particular, the percent <u>increase</u> in $P(W| A)$ = 60% relative to $P(W)$ = 32% is usually computed as

$$\left(\frac{P(W \mid A) - P(W)}{P(W)}\right) \times 100\% = \left(\frac{60\% - 32\%}{32\%}\right) \times 100\% = \left(\frac{28\%}{32\%}\right) \times 100\% = \left(\frac{7}{8}\right) \times 100\% = \mathbf{87.5\%},$$

which is precisely $= (1.875 - 1)\% \ = \mathbf{D}(W, A) - 1) \times 100\%$.

So, $P(W| A)$ = 60% is an 87.5% **increase** over $P(W)$ = 32% in the presence of the given A-classification status of the selected systems engineer.

(g) Find the conditional probability a randomly selected person is A-classified, given they are a women.

$$P(A \mid W) = \mathbf{D}(A, W) \times P(A) = \mathbf{D}(W, A) \times P(A) = \mathbf{1.875} \times 30\% = 56.25\%,$$

which we already know must be the same **87.5%** increase over $P(A)$ = 30% (Check!)

(h) Suppose a systems engineer from the IT department is randomly selected, find the probability that person is a man.

$P(M) = 0.68,$ computed directly from the second-row total in the table.

(i) Suppose a systems engineer from the IT department is randomly selected, find the probability that person is an A-classified man.

$P(M \cap A) = 0.12,$ computed directly from the second-row, first-column cell of the table.

(k) Find the coefficient of dependence factor, $\mathbf{D}(M, A)$

Using (c), (h) and (i) above,

$$\mathbf{D}(M, A) = \frac{P(M \cap A)}{P(M) \times P(A)} = \frac{0.12}{(0.68 \times 0.30)} = \frac{1{,}200}{2{,}040} = \frac{\mathbf{10}}{\mathbf{17}} \approx \mathbf{0.5882}.$$

(l) Find the conditional probability a randomly selected person is a man, given they are A-classified.

$P(M \mid A) = 1 - P(W \mid A) = 1 - 60\% = 40\%$ or, if one chooses to use (f) above,

$$P(M \mid A) = \mathbf{D}(M, A) \times P(M) = \left(\frac{\mathbf{10}}{\mathbf{17}}\right) \times 68\% = 40\%.$$

Events M and A are sub-dependent. Notice that $P(M)$ = 68% almost halves to $P(M \mid A)$ = 40% in the presence of the given information. Again, it is the coefficient of dependence factor, **0.5882 (=10/17)** that provides the precise transformational effect that the conditional information conveys beyond the marginal probability.

The percent <u>decrease</u> in $P(M| A)$ = 40% relative to $P(M)$ = 68% is computed as

$$\left(\frac{P(M \mid A) - P(M)}{P(M)}\right)\times 100\% = \left(\frac{40\% - 68\%}{68\%}\right)\times 100\% = \left(\frac{-28\%}{68\%}\right)\times 100\% = \left({}^{10}\!/_{17} - 1\right)\times 100\%$$

$$= \left({}^{-7}\!/_{17}\right)\times 100\% \approx \textbf{-41.176\%} \equiv \left(\mathbf{D}(M, A) - 1\right)\times 100\%.$$

(m) Find the conditional probability a randomly selected person is A-classified, given they are a man.

$$P(A \mid M) = \mathbf{D}(A, M)\times P(A) = \mathbf{D}(M, A)\times P(A) = \left(\mathbf{{}^{10}\!/_{17}}\right)\times 30\% \approx 17.647\%,$$

which we already know must be the same **41.176%** decrease relative to $P(A)$ = 30% seen in bullet item (l) above.

∎

A consequence of bullet items (f) and (l) is the following proposition.

Proposition 2.1: Let $\boldsymbol{S}$ be the sample space of a random experiment having probability function P defined for all events associated with $\boldsymbol{S}$. Let A and B be any two non-null events. Then, the percent change in $P(A \mid B)$ relative to $P(A)$ is computed as

$$\boxed{\%\ change\ in\ P(A \mid B)\ relative\ to\ P(A)} = \left[\mathbf{D}(A, B) - 1\right]\times 100\%,$$

where $\mathbf{D}(A,B)$ is the coefficient of dependence between events A and B. If this computation results in a positive value, then it is regarded as a percent increase, while if negative, then it is deemed a percent decrease.

Proof:

$$\begin{aligned}\%\ change &= \left[\frac{P(A \mid B) - P(A)}{P(A)}\right]\times 100\% = \left[\frac{P(A \mid B)}{P(A)} - 1\right]\times 100\% \\ &= \left[\frac{P(A \mid B)\times P(B)}{P(A)\times P(B)} - 1\right]\times 100\% \\ &= \left[\frac{P(A \cap B)}{P(A)\times P(B)} - 1\right]\times 100\% \\ &= \left[\mathbf{D}(A, B) - 1\right]\times 100\%.\end{aligned}$$

∎

The next theorem collects some results, whose proofs appear in the Appendix for interested readers. More advanced students can be challenged to demonstrate these claims.

Theorem 2.1: Let A and B be any two non-null events and let A^c denote the complement of A. Then, if $\mathbf{D}(A, B)$ denotes the coefficient of dependence between A and B, then

a) $\mathbf{D}(A,B) = 0$ if and only if $P(A \cap B) = 0$.
b) If A and B are disjoint events $(A \cap B = \varnothing)$, then $\mathbf{D}(A,B) = 0$.
c) $\mathbf{D}(A,B) > 1$ if and only if $\mathbf{D}(A^c,B) < 1$.
d) $\mathbf{D}(A,B) = 1$ if and only if $\mathbf{D}(A^c,B) = 1$.
e) $\mathbf{D}(A,B) < 1$ if and only if $\mathbf{D}(A^c,B) > 1$.
f) $\mathbf{D}(A^c,B) = \dfrac{1-\mathbf{D}(A,B)\times P(A)}{\left(1-P(A)\right)} = \mathbf{D}(B,A^c)$.
g) $\mathbf{D}(A,A) = \dfrac{1}{P(A)} \geq 1$.

Proof: See Appendix. ∎

Corollary 2.1: Let A and B be any two non-null events and let $\mathbf{D}(A, B)$ denote the coefficient of dependence between A and B, then

a) $\mathbf{D}(A,B) > 1$ **if and only if** $P(A \mid B) > P(A)$ **if and only if** $P(B \mid A) > P(B)$.
b) $\mathbf{D}(A,B) = 1$ **if and only if** $P(A \mid B) = P(A)$ **if and only if** $P(B \mid A) = P(B)$.
c) $\mathbf{D}(A,B) < 1$ **if and only if** $P(A \mid B) < P(A)$ **if and only if** $P(B \mid A) < P(B)$.

The proof for Corollary 2.1 is similar to those given for the proof of Theorem 2.1, which appears in the Appendix ∎

The converse of Theorem 2.1(b) does not hold. That is, $\mathbf{D}(A,B) = 0$ **does not guarantee** that A and B are disjoint events, as the next example demonstrates.

Example 2.2: (Uniform distribution). The experiment here is to choose a number "at random" from the unit interval [0,1]. Suppose $A = \left\{x : 0 \leq x \leq \frac{1}{2}\right\}$ and $B = \left\{x : \frac{1}{2} \leq x \leq 1\right\}$, then $A \cap B = \left\{x = \frac{1}{2}\right\}$. Moreover, $P(A) = \dfrac{0.5-0}{1.0-0} = 0.5$, $P(B) = \dfrac{1.0-0.5}{1.0-0} = 0.5$, and

$$P(A \cap B) = P\left(\frac{1}{2} \leq x \leq \frac{1}{2}\right) = \frac{0.5-0.5}{1.0-0} = 0.$$

So, $\mathbf{D}(A,B) = \dfrac{P(A \cap B)}{P(A)\times P(B)} = \dfrac{0}{0.5\times 0.5} = 0$, even though events A and B are not disjoint. ∎

The next example nuances the conclusion of Theorem 2.1(g) to two different non-null events *A* and *B* by showing that the range of **D**(*A*,*B*) is actually $\{d : d \geq 0\}$.

Example 2.3: (Exponential distribution). The random experiment here comes from the unit exponential distribution which has distribution function

$$P(X \leq x) = F(x) = 1 - \exp(-x) \text{ when } x \geq 0.$$

Choose any positive value, say *w*. Let $A = \{x : x > w+1\}$ and $B = \{x : x > w\}$. Then, $A \cap B = \{x : x > w+1\} \cap \{x : x > w\} = \{x : x > w+1\}$, so that

$$P(A) = P(\{x : x > w+1\}) = 1 = P(\{x : x \leq w+1\}) = 1 - F(w+1) = 1 - (1 - \exp(-(w+1))) = \exp(-(w+1))$$
$$P(B) = P(\{x : x > w\}) = 1 = P(\{x : x \leq w\}) = 1 - F(w) = 1 - (1 - \exp(-w)) = \exp(-w)$$
$$P(A \cap B) = \exp(-(w+1)).$$

Now, letting $w \to +\infty$ shows

$$\mathbf{D}(A,B) = \frac{P(A \cap B)}{P(A) \times P(B)} = \frac{\exp(-(w+1))}{(\exp(-(w+1))) \times (\exp(-w))} = \frac{1}{\exp(-w)} = \exp(w) \to +\infty,$$

demonstrating the range of $\mathbf{D}(A,B)$ to be the set of non-negative, real numbers; $0 \leq \mathbf{D}(A,B) < +\infty.$ ∎

Theorem 2.2 provides a consistency check on calculations involving the coefficient of dependence, and its proof is provided in the Appendix as well.

Theorem 2.2: Let *A* be a non-null event and suppose $\{B_1, B_2, ..., B_n\}$ is a finite partition (pairwise mutually exclusive events with $B_1 \cup B_2 \cup ...... \cup B_n = \mathbf{S}$) of a sample space. Then,

$$1 = \mathbf{D}(A,B_1) \times P(B_1) + \mathbf{D}(A,B_2) \times P(B_2) + ... + \mathbf{D}(A,B_n) \times P(B_n)$$
$$= \sum_{i=1}^{n} \mathbf{D}(A,B_i) \times P(A) \times P(B_i).$$

∎

Example 2.4: Revisit Example 2.1 wherein

$$\mathbf{D}(A,W) = \frac{15}{8}, \mathbf{D}(A,M) = \frac{10}{17}, P(W) = 0.32, \text{ and } P(M) = 0.68.$$

Then, a quick consistency check is

$$\mathbf{D}(A,W) \times P(W) + \mathbf{D}(A,M) \times P(M) = \left(\frac{15}{8}\right) \times 0.32 + \left(\frac{10}{17}\right) \times 0.68 = 1.00.$$

3. Conclusions

The coefficient of dependence is introduced and studied here as a means of helping students truly understand the elusive topics of independence versus dependence. Additionally, several illustrative examples are given to help instructors in their quest to both broaden and strengthen student understanding of conditional probability via the coefficient of dependence. We have long held the belief that the best explanation for any concept is the simplest one that gives the essential insight, and the coefficient of dependence is precisely the transformation to teach for connecting marginal and conditional probabilities.

References

[1]. Freund, J. (2004). *Modern Elementary Statistics* (11th ed.). Upper Saddle River, NJ. Pearson. Prentiss-Hall.
[2]. Hogg, R. V., Tanis, E. A. and Zimmerman, D. L (2020). *Probability and Statistical Inference* (10th ed.) Hoboken, NJ. Pearson Education.
[3]. Navidi, W. Monk, B. (2022). *Elementary Statistics* (4th ed.). New York, NY. McGraw Hill.
[4]. Ott, R. L. and Longnecker, M. (2010). *An Introduction to Statistical Methods and Data Analysis* (6th ed.). Belmont, CA. Brooks/Cole

APPENDIX

Proof for Theorem 2.1:

a) $$\mathbf{D}(A,B)=0 \Leftrightarrow \frac{P(A\cap B)}{P(A)\times P(B)}=0 \Leftrightarrow P(A\cap B)=0.$$

b) $$A\cap B=\varnothing \Rightarrow \mathbf{D}(A,B)=\frac{P(A\cap B)}{P(A)\times P(B)}=\frac{P(\varnothing)}{P(A)\times P(B)}=\frac{0}{P(A)\times P(B)}=0.$$

c) $$\mathbf{D}(A,B)>1 \Leftrightarrow \frac{P(A\cap B)}{P(A)\times P(B)}>1 \Leftrightarrow \frac{P(A\cap B)}{P(B)}>P(A) \Leftrightarrow P(A\,|\,B)>P(A) \Leftrightarrow 1-P(A\,|\,B)<1-P(A)$$
$$\Leftrightarrow P(A^c\,|\,B)<P(A^c) \Leftrightarrow \frac{P(A^c\cap B)}{P(B)}<P(A^c) \Leftrightarrow \frac{P(A^c\cap B)}{P(A^c)\times P(B)}<1 \Leftrightarrow \mathbf{D}(A^c,B)<1.$$

d) $$\mathbf{D}(A,B)=1 \Leftrightarrow \frac{P(A\cap B)}{P(A)\times P(B)}=1 \Leftrightarrow P(A\cap B)=P(A)\times P(B) \Leftrightarrow \left[P(B)-P(A^c\cap B)\right]=P(A)\times P(B)$$
$$\Leftrightarrow \left[P(A^c\cap B)-P(B)\right]=-P(A)\times P(B) \Leftrightarrow P(A^c\cap B)=P(B)-P(A)\times P(B) \Leftrightarrow$$
$$P(A^c\cap B)=\left(1-P(A)\right)\times P(B) \Leftrightarrow P(A^c\cap B)=P(A^c)\times P(B) \Leftrightarrow \frac{P(A^c\cap B)}{P(A^c)\times P(B)}=1 \Leftrightarrow \mathbf{D}(A^c,B)=1$$

e)

$$\mathbf{D}(A,B)<1 \;\Leftrightarrow\; \frac{P(A\cap B)}{P(A)\times P(B)}<1 \Leftrightarrow \frac{P(A\cap B)}{P(B)}<P(A) \Leftrightarrow P(A\,|\,B)<P(A) \Leftrightarrow 1-P(A\,|\,B)>1-P(A)$$

$$\Leftrightarrow P(A^c\,|\,B)>P(A^c) \Leftrightarrow \frac{P(A^c\cap B)}{P(B)}>P(A^c) \Leftrightarrow \frac{P(A^c\cap B)}{P(A^c)\times P(B)}>1 \Leftrightarrow \mathbf{D}(A^c,B)>1.$$

f)

$$\mathbf{D}(A^c,B)=\frac{P(A^c\cap B)}{P(A^c)\times P(B)}=\frac{P(B)-P(A\cap B)}{P(A^c)\times P(B)}=\left[\frac{P(B)}{P(A^c)\times P(B)}\right]-\left[\frac{P(A\cap B)}{P(A^c)\times P(B)}\right]\left(\frac{P(A)}{P(A)}\right)$$

$$=\left[\frac{1}{1-P(A)}\right]-\left[\frac{P(A\cap B)}{[1-P(A)]\times P(B)}\right]\left(\frac{P(A)}{P(A)}\right)=\left[\frac{1}{1-P(A)}\right]-\left[\frac{1}{1-P(A)}\right]-\left[\frac{P(A\cap B)}{P(A)\times P(B)}\right]\times P(A)$$

$$=\frac{1-\mathbf{D}(A^c,B)\times P(A)}{1-P(A)}.$$

g) $$\mathbf{D}(A,A)=\frac{P(A\cap A)}{P(A)\times P(A)}=\frac{P(A)}{P(A)\times P(A)}=\frac{1}{P(A)}\geq 1.$$

∎

<u>Proof for Theorem 2.2</u>:

$$\sum_{i=1}^{n}\mathbf{D}(A,B_i)\times P(A)\times P(B_i)=\sum_{i=1}^{n}\left[\frac{P(A\cap B_i)}{P(A)\times P(B_i)}\right]\times P(A)\times P(B_i)=\sum_{i=1}^{n}\left[\frac{P(A)\times P(B_i)}{P(A)\times P(B_i)}\right]\times P(A\cap B_i)$$

$$=\sum_{i=1}^{n}(1\times P(A\cap B_i)=P(A).$$

∎