1 Cramér’s Contingency Coefficient

The most common used measure is

\[
\text{Cramér's coefficient} = \sqrt{R_1} = \sqrt{\frac{T}{N(q - 1)}}
\]

where \( q = \min\{r, c\} \).

2 Pearson’s Contingency Coefficient

Two other coefficient are sometimes used. The first is called \textit{Pearson's coefficient of mean square contingency} by Yule and Kendall and is given as

\[
R_2 = \sqrt{\frac{T}{N + T}}
\]

\( R_2 \) is also called the \textit{contingency coefficient} by McNemar and Siegel.

3 Pearson’s Mean-Square Contingency Coefficient

We present a third measure of dependence, \( R_3 \), also attributed to Pearson and also called the \textit{mean-square contingency}. \( R_3 \) is defined as

\[
R_3 = \frac{T}{N}
\]

Finally, we just mention \textit{Tschuprow's coefficient}, given by Yule and Kendall, as

\[
R_4 = \sqrt{\frac{T}{N\sqrt{(r - 1)(c - 1)}}}
\]

4 The Phi Coefficient

If the type of association is of interest, care must be taken to set up the table so that \( a \) and \( d \) represent the number of similar classifications (dark-dark and light-light), while \( b \) and \( c \) represent the number of unlike classifications (dark-light and light-dark). [Regarding a 2×2 contingency table] One measure of association that preserves direction [since \( R_1-R_4 \) are all positive] is the \textit{phi coefficient}, given by

\[
R_5 = \frac{ad - bc}{\sqrt{r_1r_2c_1c_2}}
\]

The phi coefficient is merely Cramér’s coefficient with the sign of \( (ad - bc) \) being preserved.
Other measures of association for the four-fold contingency table include one proposed by Yule and Kendall

\[ R_6 = \frac{ad - bc}{ad + bc} \]  

and one proposed by Ives and Gibbons

\[ R_7 = \frac{(a + d) - (b + c)}{a + b + c + d} \]

There is no end to the possible measures that may be defined. One’s choice of a coefficient is solely a result of personal preferences.

5 Exercises 4.4 #1, 3

I used R to solve the following problems, click here to view source code.

1. One hundred married couples were interviewed, and the husband and wife were asked separately for their first choice for the next U.S. president, with the following results. [Look at page 238]
   (a) T = 6.34  (b) \( \sqrt{R_1} = 0.1780449 \)  (c) \( R_1 = 0.0317 \)  (d) \( R_2 = 0.2441722 \)  (e) \( R_3 = 0.0634 \)  (f) \( R_4 = 0.1780449 \)

3. A traffic study was conducted for a short time on a well-traveled city street. Of the 64 cars observed, 16 were exceeding the limit and 48 were not. Also, 24 of the drivers had passengers and the rest did not. Twelve of the speeders were driving alone. Assume that the observed traffic behaves the same as a random sample of all traffic would.
   (a) \( R_5 = 0.1490712 \), the \( p \)-value is .116519.  (b) \( R_6 = 0.3636364 \)  (c) \( R_7 = 0 \)