The last model that we consider in discussing the theoretical approach to CSC construction is the one proposed by D. O. Avetisyan (1975, 1977), which leads to a complex search characteristic we have not mentioned. Two random variables are used within Avetisyan's model: "to be an output document" and "to be a pertinent document." These variables will be denoted by X and Y, respectively, and could have the following values for a specific document: $X = 1$ if the document was found during search and $X = 0$ otherwise; $Y = 1$ if the document is pertinent and $Y = 0$ otherwise. It is clear that the "stronger" the connection between the two variables, the "closer" the produced output to the "ideal," that is, the better the outcome of the search. In other words, in an evaluation of the search results, one can attempt to use a tool that makes it possible to determine if there is a direct dependence between variables X and Y and what the "strength" of this dependence is.

The discussed model has one noteworthy feature: the model does not represent actual results of the search. At the same time it allows one to evaluate the search results, a feature quite adequate for the construction of complex search characteristics.

To determine if there exists a dependence between variables X and Y and what the "strength" of this dependence is, the model employs a coefficient of linear correlation

$$\frac{E(XY) - E(X) \cdot E(Y)}{\sqrt{D(X) \cdot D(Y)}},$$

where $E(XY)$, $E(X)$, and $E(Y)$ are expected values of random variables XY, X and Y, and $D(X)$, and $D(Y)$ are a dispersions of variables X and Y.

Using approaches common for such cases for determining the values of the components in the preceding expression, it was established within the framework of the model under consideration that

$$\frac{E(XY) - E(X) \cdot E(Y)}{\sqrt{D(X) \cdot D(Y)}} = \frac{rd - lb}{\sqrt{(r + l)(r + b)(b + d)(l + d)}}.$$

(Intermediate calculations are omitted because they are both clear and tedious.) Thus, the discussed model gives us a new complex search characteristic:

$$I_{12} = \frac{rd - lb}{\sqrt{(r + l)(r + b)(b + d)(l + d)}}.$$

The new CSC differs from the CSCs discussed earlier. However, under certain conditions the values of this new CSC will be close to those for CSC $I_2 = \sqrt{R \cdot P}$. This follows from the expression known from the folklore: $\lim_{d \to \infty} I_{12} = I_2$ (assuming that r, b, and l are bounded). We will demonstrate that this correlation does exist (regretfully, the authors failed to locate any publications of this correlation and its proof):

$$\lim_{d \to \infty} I_{12} = \lim_{d \to \infty} \frac{rd - lb}{\sqrt{(r + l)(r + b)(b + d)(l + d)}}$$

$$= \lim_{d \to \infty} \frac{r - \dfrac{lb}{d}}{\sqrt{(r + l)(r + b)\left(\dfrac{b}{d} + 1\right)\left(\dfrac{l}{d} + 1\right)}}$$

$$= \frac{r}{\sqrt{(r + l)(r + b)}} = \frac{r}{\sqrt{NC}} = \sqrt{R \cdot P} = I_2.$$

Thus, if d is much larger than $l \cdot b$, the values of characteristic I_{12} are indeed very close to characteristic I_2 values. Nevertheless, the complex search characteristics I_2 and I_{12} are substantially different, as will be seen from our further discussion (see Section 10.8, "Order Preservation Property").

Note that for complex search characteristic I_{12}, one can also derive the more common appearance (Avetisyan, 1975), shown here:

$$I_{12} = \frac{rd - lb}{\sqrt{(r + l)(r + b)(b + d)(l + d)}} = \frac{rd - lb}{\sqrt{N \cdot C \cdot (b + d) \cdot L}}$$

$$= \frac{\sqrt{r \cdot l} \cdot (rd - lb)}{\sqrt{N \cdot C \cdot L \cdot (r \cdot l \cdot b + r \cdot l \cdot d)}}$$

$$= \frac{\dfrac{\sqrt{r \cdot l} \cdot [rd - (L - d)(C - r)]}{N \cdot C \cdot L}}{\sqrt{\dfrac{r \cdot l \cdot b + r \cdot l \cdot d + r^2 \cdot d - r^2 \cdot d}{N \cdot C \cdot L}}}$$

$$= \frac{\sqrt{\dfrac{r \cdot l}{N^2}} \cdot \dfrac{rd - LC + dC + rL - rd}{C \cdot L}}{\sqrt{\dfrac{rd(l + r) - r(rd - lb)}{N \cdot C \cdot L}}} = \frac{\sqrt{\dfrac{r(N - r)}{N \cdot N}} \cdot \left(\dfrac{r}{C} + \dfrac{d}{L} - 1\right)}{\sqrt{\dfrac{r}{C} \cdot \dfrac{d}{L} \cdot \dfrac{N}{N} - \dfrac{r}{N} \cdot \dfrac{rd - lb}{C \cdot L}}}$$

$$= \frac{\sqrt{P(1 - P)} \cdot (R + S - 1)}{\sqrt{R \cdot S - P(R + S - 1)}}.$$

Finally, it is noteworthy that the linear correlation coefficient does not seem to be the only tool that will enable one to determine whether there is a direct dependence between random variables X and Y and what the "strength"