probabilities), although the logistic function can

ϕ has a value of 0 when there is no association between the two values, that is, when the probability of selecting a value of ***A*** is constant for any value of ***B***.

2. **n n nc** In Appendix 1 we prove that Cramér's $\phi(\boldsymbol{A}, \boldsymbol{B})$ ϕ measures the linear interpolation from flat to identity matrix (see Figure 1 for the idea). We can therefore refer to ϕ as the best estimate of the population **n n n o y** ***idp*** $\equiv$ ***p***(***a*** $\leftrightarrow$ ***b***). This intercorrelation is robust, i.e. it 'gracefully decays' the further it deviates from this ideal. ϕ is

$$\boldsymbol{dp_R}(\boldsymbol{A}, \boldsymbol{b}) \equiv -\sum_{i=}^{k} {}_R\ \boldsymbol{a}_i\ \boldsymbol{b}$$

We note the following.

1. Relative dependency $dp_R(B, A)$ is linear with x when $p(a)$ is even.

2. Measures are ordered in size: $C_{adj} > dp_R(B, A) > \phi > dp_R(A, B)$.

3. $dp_R(A, B)$ (and therefore ϕ) converges to $dp_R(B, A)$ as $p(b)$ becomes more even (tends to $1/k$).

Whereas dp_R measures the distance on the first parameter from the prior (and is thus directional when a prior skew is applied to one variable only), ϕ is based on the root mean square distance of both variables. C_{adj} appears to behave rather differently to ϕ, as the right hand graph in Figure 2 shows. Given that the only other bi-directional measure, ϕ, measures the interdependence of A and B, there appears to be little advantage in adopting the less conservative C_{adj}.

Finally, for $k = 2$ the following equation also holds:

4. $\phi^2 = dp_R(A, B) \times dp_R(B, A)$.

We find an equality between a classical Bayesian approach to dependency and a stochastic approach based on Pearson's χ^2 for one degree of freedom. The proof is given in Appendix 3.

This raises the following question: what does 'directionality' mean here?

Note that dp_R

6. A worked example

Figure 3 provides a demonstration of plotting confidence intervals on ϕ

Appendix 1. The best estimate of population interdependent probability is Cramér's ϕ

Cramér's ϕ is not merely a 'measure of association'. It repres

Appendix 2. Deriving 2 × 2 rule dependency

For a 2×2 table with a single degree of freedom, the following axioms hold.

A1. $\boldsymbol{p}(\boldsymbol{a}_2) = \boldsymbol{p}(\neg\boldsymbol{a}_1) = 1 - \boldsymbol{p}(\boldsymbol{a}_1)$; $\boldsymbol{p}(\boldsymbol{a}_2 \mid \boldsymbol{b}_i) = \boldsymbol{p}(\neg\boldsymbol{a}_1 \mid \boldsymbol{b}_i) = 1 - \boldsymbol{p}(\boldsymbol{a}_1 \mid \boldsymbol{b}_i)$,

A2. $\boldsymbol{p}(\boldsymbol{a}_1 \mid \boldsymbol{b}_i) - \boldsymbol{p}(\boldsymbol{a}_1) = \boldsymbol{p}(\boldsymbol{a}_2) - \boldsymbol{p}(\boldsymbol{a}_2 \mid \boldsymbol{b}_i)$,

A3. $[\boldsymbol{p}(\boldsymbol{a}_1 \mid \boldsymbol{b}_i) < \boldsymbol{p}(\boldsymbol{a}_1)] \leftrightarrow [\boldsymbol{p}(\boldsymbol{a}_2) < \boldsymbol{p}(\boldsymbol{a}_2 \mid \boldsymbol{b}_i)]$.

A1 is a consequence of the Boolean definition of $\boldsymbol{A}$, A2 can be demonstrated using Bayes' Theorem and A3 is a consequence of A2. A2 further entails that row sums are equal, i.e. $\boldsymbol{dp_R}(\boldsymbol{a}_1, \boldsymbol{B}) = \boldsymbol{dp_R}(\boldsymbol{a}_2, \boldsymbol{B})$.

Equation (5) may therefore be simplified as follows

$$\boldsymbol{dp_R}(\boldsymbol{A}, \boldsymbol{B}) \equiv \frac{1}{\boldsymbol{k}} \sum_{\boldsymbol{i}=1}^{\boldsymbol{k}} \sum_{\boldsymbol{j}=1}^{\boldsymbol{k}} \boldsymbol{dp_R}(\boldsymbol{a}_i, \boldsymbol{b}_j) \times \boldsymbol{p}(\boldsymbol{b}_j) = \sum_{\boldsymbol{j}=1}^{\boldsymbol{k}} \boldsymbol{dp_R}(\boldsymbol{a}_1, \boldsymbol{b}_j) \times \boldsymbol{p}(\boldsymbol{b}_j)$$
$$= \frac{\boldsymbol{p}(\boldsymbol{a}_1 \mid \boldsymbol{b}_1) - \boldsymbol{p}(\boldsymbol{a}_1)}{1 - \boldsymbol{p}(\boldsymbol{a}_1)} \times \boldsymbol{p}(\boldsymbol{b}_1) + \frac{\boldsymbol{p}(\boldsymbol{a}_1) - \boldsymbol{p}(\boldsymbol{a}_1 \mid \boldsymbol{b}_2)}{\boldsymbol{p}(\boldsymbol{a}_1)} \times \boldsymbol{p}(\boldsymbol{b}_2).$$

Applying Bayes' Theorem ($\boldsymbol{p}(\boldsymbol{a}_1 \mid \boldsymbol{b}_2) \equiv \boldsymbol{p}(\boldsymbol{b}_2 \mid \boldsymbol{a}_1) \times \boldsymbol{p}(\boldsymbol{a}_1) / \boldsymbol{p}(\boldsymbol{b}_2)$) and axiom A1:

$$\boldsymbol{dp_R}(\boldsymbol{A}, \boldsymbol{B}) = \frac{\boldsymbol{p}(\boldsymbol{a}_1 \mid \boldsymbol{b}_1)\boldsymbol{p}(\boldsymbol{b}_1)}{1 - \boldsymbol{p}(\boldsymbol{a}_1)} - \frac{\boldsymbol{p}(\boldsymbol{a}_1)\boldsymbol{p}(\boldsymbol{b}_1)}{1 - \boldsymbol{p}(\boldsymbol{a}_1)} + \frac{\boldsymbol{p}(\boldsymbol{a}_1 \mid \boldsymbol{b}_1)\boldsymbol{p}(\boldsymbol{b}_1)}{\boldsymbol{p}(\boldsymbol{a}_1)} - \boldsymbol{p}(\boldsymbol{b}_1).$$

The first and third terms then simplify to $[\boldsymbol{p}(\boldsymbol{a}_1 \mid \boldsymbol{b}_1)\boldsymbol{p}(\boldsymbol{b}_1)] / [(1 - \boldsymbol{p}(\boldsymbol{a}_1))\boldsymbol{p}(\boldsymbol{a}_1)]$, so

$$\boldsymbol{dp_R}(\boldsymbol{A}, \boldsymbol{B}) = \frac{\boldsymbol{p}(\boldsymbol{a}_1 \mid \boldsymbol{b}_1)\boldsymbol{p}(\boldsymbol{b}_1) - \boldsymbol{p}(\boldsymbol{a}_1)^2\boldsymbol{p}(\boldsymbol{b}_1) - (1 - \boldsymbol{p}(\boldsymbol{a}_1))\boldsymbol{p}(\boldsymbol{a}_1)\boldsymbol{p}(\boldsymbol{b}_1)}{(1 - \boldsymbol{p}(\boldsymbol{a}_1))\boldsymbol{p}(\boldsymbol{a}_1)}$$
$$= \frac{[\boldsymbol{p}(\boldsymbol{a}_1 \mid \boldsymbol{b}_1) - \boldsymbol{p}(\boldsymbol{a}_1)]\boldsymbol{p}(\boldsymbol{b}_1)}{[1 - \boldsymbol{p}(\boldsymbol{a}_1)]\boldsymbol{p}(\boldsymbol{a}_1)}. \qquad \square$$

Appendix 3. For a 2 × 2 table, $\phi^2 \equiv dp_R(A, B) \times dp_R(B, A)$

The proof is in three stages.

S

STAGE 2. Converting to a+b+c+d notation.

The 2 × 2 χ^2 statistic, and thus ϕ, may be represented simply in terms of four frequencies in the table, a, b, c and d (note roman font to distinguish from $\boldsymbol{a}$, $\boldsymbol{a}_1$, etc). The table is labelled thus, and $\boldsymbol{N} \equiv$ a+b+c+d:

	$\boldsymbol{b}_1$	$\boldsymbol{b}_2$	Σ
$\boldsymbol{a}$			