n 人を対象として，2 つの質問（X，Y）をした（それぞれの質問を変数 X，Y とする)。
2 つの質問はそれぞれ「はい」か「いいえ」のいずれかを回答するものとする。
得られたデータを分析するときに，各対象者がそれぞれの質問に「いいえ」と答えている場合は 0,「はい」と答えている場合は 1 を対応させる。
例えば，k番目の対象が質問 X に「はい」，質問 Y に「いいえ」と答えているときには，$X_k = 1$，$Y_k = 0$ となる。表 1 ではこのようなデータを $(1,0)$ と表している。
調査の結果得られたデータ $(X_i, Y_i), i = 1, 2, \cdots, n$ をまとめると，表 1 のようになったとする。

表 1　2 変数データ (X_i, Y_i)

(X_i, Y_i)	
$(0,0)$ $\vdots$ $(0,0)$	質問 X への回答が「いいえ」で，かつ質問 Y への回答が「いいえ」のものが a 人
$(0,1)$ $\vdots$ $(0,1)$	質問 X への回答が「いいえ」で，かつ質問 Y への回答が「はい」のものが b 人
$(1,0)$ $\vdots$ $(1,0)$	質問 X への回答が「はい」で，かつ質問 Y への回答が「いいえ」のものが c 人
$(1,1)$ $\vdots$ $(1,1)$	質問 X への回答が「はい」で，かつ質問 Y への回答が「はい」のものが d 人

このとき，変数 X と変数 Y のピアソンの積率相関係数は，(1) 式で定義できる。

$$(1)\quad r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\ \sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \quad \text{ただし，} \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

一方，得られたデータを 2×2 分割表として集計すると，表 2 のようになる。

表 2　2×2 分割表における記号の定義

	質問 Y への回答		
質問 X への回答	いいえ	はい	合計
いいえ	a	b	$e(= a + b)$
はい	c	d	$f(= c + d)$
合計	$g(= a + c)$	$h(= b + d)$	$n(= a + b + c + d)$

表 2 の分割表において，質問 X と質問 Y の間の属性相関係数 ϕ は (2) 式で与えられる。

$$(2)\quad \phi = \frac{|\,a\,d - b\,c\,|}{\sqrt{e\,f\,g\,h}}$$

質問 1： 変数 X の平均値 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ を，表 2 の記号（a，b，$\cdots$，h，n）を使って表せ。

質問 2： 変数 X の変動 $\sum_{i=1}^{n}(X_i - \bar{X})^2$ を表 2 の記号（a，b，$\cdots$，h，n）を使って表せ。

質問 3：（1）式で得られるピアソンの積率相関係数 r と（2）式で得られる属性相関係数 ϕ において，$\phi = |r|$ となることを示せ。