Correlation of two random variables

The correlation ρ of X and Y is defined as

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}}$$

the denominator is like a normalizing factor, i.e. balances the size of the numerator.

Fact: $-1 \le \rho(X,Y) \le +1$ for all random variables X, Y.

Example: Go back to the hat problem,
let X indicate if Alice gets her hat back, Y indicates if Bob gets his back.

Know $Cov(X,Y) = \frac{41}{225}$.

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{41/225}{\sqrt{(\frac{1}{10})(\frac{9}{10})(\frac{1}{10})(\frac{9}{10})}} = \frac{1}{81} = 0.0123$$

Idea: If the correlation is near 1, then X large $\Rightarrow$ Y is somewhat larger,
is near -1, then X large $\Rightarrow$ Y is somewhat smaller,

(i.e. near $+1$, Y decreases as X decreases or increases as X increases i.e. they correlated in the same direction)

(near -1, Y decreases as X increases or vice versa)

If ρ is exactly 0, we say X and Y are uncorrelated.
(Different from independence.)
I.e. if X increases, and ρ near 0, doesn't tell us much about what happens to Y.