Let X be a (real-valued) random variable described by a cumulative distribution function (cdf) $F_X(x) = \Pr(X \leq x)$, which in turn is described either by a probability density function (pdf) $f_X(x) = dF_X(x)/dx$ if X is continuous, or a probability mass function (pmf) $p_X(x)$ if X is discrete. Let Y be another random variable with cdf F_Y etc. A joint cdf for both X and Y is denoted by $F_{XY}(x, y) = \Pr(X \leq x, Y \leq y)$.

Assume throughout that $E(X) = E(Y) = 0$,
$E(X^2) = \sigma_X^2$, $E(Y^2) = \sigma_Y^2$.
Both σ_X^2 and σ_Y^2 are assumed to be nonzero and finite.

A very old and very useful measure of "distance" between two given cdfs F_X and F_Y is defined by

$$\overline{d}(F_X, F_Y) = \min_{F_{XY}} E[(X - Y)^2],$$

where the expectation is with respect to the joint cdf F_{XY} and the minimum is over all joint cdfs F_{XY} having the given F_X and F_Y as marginals.

First Question: Given arbitrary cdfs F_X and F_Y describing random variables X and Y, give a *simple* example of a joint cdf F_{XY} with the prescribed marginals and use it to find an upper bound to $\overline{d}(F_X, F_Y)$ which depends only on σ_X^2 and σ_Y^2.

Solution: Assume that X and Y are independent random variables, in which case $F_{XY}(x, y) = F_X(x)F_Y(y)$ and $E(XY) = E(X)E(Y) = 0$ and hence

$$E[(X - Y)^2] = E(X^2) + E(Y^2) - 2E(XY) = \sigma_X^2 + \sigma_Y^2.$$