Third Question: Suppose that both X and Y are Gaussian (0 means, variances σ_X^2 and σ_Y^2 respectively).

Find $\overline{d}(F_X, F_Y)$.

Solution: In a short oral exam there is no time for formal optimization. The idea here is to realize (possibly with a hint) that if you can think of a joint distribution with the given marginals which hits the previous lower bound, then that must be the minimum since no joint distribution can do any better. Here there is a natural guess — you can turn a 0 mean Gaussian X with variance σ_X^2 into a 0 mean Gaussian Y with variance σ_Y^2 by a simple scaling $Y = \sigma_Y X/\sigma_X$, which yields

$$\begin{aligned} E[(X-Y)^2] &= \sigma_X^2 + \sigma_Y^2 - 2E(XY) \\ &= \sigma_X^2 + \sigma_Y^2 - 2\frac{\sigma_Y}{\sigma_X}E(X^2) \\ &= \sigma_X^2 + \sigma_Y^2 - 2\sigma_X\sigma_Y = (\sigma_X - \sigma_Y)^2. \end{aligned}$$

Since this is an unbeatable lower bound from the previous part, it must solve the minimization. Note that here one of the two random variables is defined as a deterministic random variable, which is how Monge defined his distance over two centuries ago. The definition in terms of fixed marginals and a best joint was Kantorovich's. A great deal of research has been done to determine when the two definitions are equivalent.