Suppose that a scientist is planning a clinical trial with one control condition and two different treatments. She intends to simultaneously test the mean response to Treatment 1 versus the Control and Treatment 2 versus the Control; both comparisons are of equal interest.

She is ready to assume an entirely traditional univariate linear model with normal error terms. That is, $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{X}$ is an n by p matrix of known constants (for convenience, assume the rank of $\mathbf{X}$ to be p), $\boldsymbol{\beta}$ is a p by 1 matrix of unknown constants, and $\boldsymbol{\epsilon}$ is a n by 1 multivariate normal random vector with mean zero and dispersion matrix $\sigma^2\mathbf{I}_n$; σ^2 is an unknown constant.

1. Describe the columns of the $\mathbf{X}$ matrix for the problem at hand. There are uncountably many correct answers, but some are more convenient than others for the final part of this question.

2. The question of interest to the scientist needs to be expressed in terms of a null hypothesis of the form $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{h}$, where $\mathbf{C}$ is a d by p matrix of rank $d \leq p$. Write down the $\mathbf{C}$, $\boldsymbol{\beta}$ and $\mathbf{h}$ matrices.

3. Under this model, the usual test statistic has a non-central F distribution with degrees of freedom d and $n - p$, and non-centrality parameter

$$\delta^2 = \frac{(\mathbf{C}\boldsymbol{\beta} - \mathbf{h})'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}(\mathbf{C}\boldsymbol{\beta} - \mathbf{h})}{\sigma^2}.$$

   Of course larger values of δ^2 lead to larger probabilities of rejecting H_0. Suppose that in reality, mean response to the two treatments is identical, and different from the mean response for the control (so the null hypothesis is false). Suppose also that the scientist has enough funding to test 150 patients. How should she allocate them to experimental conditions? That is, give three sample sizes that add to 150.