Jason Abaluck

optimal Representation

Ephraim Sutherland

Contents

1 Setup

1. Suppose a physician can only observe an ATE and some measure of representativeness. Physicians have prior $\bar{\beta}$ and $\beta_{ATE} = (1/N)\sum \beta_i$.

2. We need a model for the betas related to each other based on x's – an individuals characteristics.

   Suppose

   $$\beta(x_i) = x_i\gamma$$

   Where x_i is a vector of characteristics and γ is a vector of coefficients.
   If you know γ, then you know β for any given patient.

3. However, you don't observe γ, you instead observe: $\beta_{ATE} = \bar{x}\gamma$ where $\bar{x} = (\frac{1}{N})\sum x_i$

4. We know β_i for patients with characteristics $\bar{x}$ (it is β_{ATE}).

5. For other patients, we need to solve

   $$\beta_{i,post} = \mathrm{E}(x_i\gamma|\bar{x}\gamma = \beta_{ATE})$$

6. To solve

   (a)

   $$\begin{aligned}\beta_{i,post} &= \mathrm{E}(x_i\gamma|\bar{x}\gamma = \beta_{ATE}) \\ &= \mathrm{E}((x_i - c_i\bar{x})\gamma|\bar{x}\gamma = \beta_{ATE}) + c_i\mathrm{E}(\bar{x}\gamma|\bar{x}\gamma = \beta_{ATE}) \\ &= \mathrm{E}((x_i - c_i\bar{x})\gamma|\bar{x}\gamma = \beta_{ATE}) + c_i\beta_{ATE}\end{aligned}$$

   For any constant c_i.
   Choose c_i so that

   $$\mathrm{Cov}((x_i - c_i\bar{x})\gamma, \bar{x}\gamma) = 0$$

Assume normality so that this guarantees independence. Then,

$$\mathrm{E}((x_i - c_i\bar{x})\gamma | \bar{x}\gamma = \beta_{ATE}) = (x_i - c_i\bar{x})\mathrm{E}(\gamma)$$

So then

$$(x_i - c_i\bar{x})\mathrm{E}(\gamma) + c_i\beta_{ATE} = x_i\mathrm{E}(\gamma) + c_i(\beta_{ATE} - \bar{x}\mathrm{E}(\gamma))$$

(c_i depends on x_i)

In other words, your belief is your prior, adjusted based on the difference between the observed ATE and your prior about the ATE. The key question is how much adjustment you do which depends on "c_i". We choose c_i to solve:

$$\mathrm{Cov}((x_i - c_i\bar{x})\gamma, \bar{x}\gamma) = 0$$

$$\iff \mathrm{Cov}(x_i\gamma, \bar{x}\gamma) - c_i\mathrm{Cov}(\bar{x}\gamma, \bar{x}\gamma) = 0$$

$$\iff \mathrm{Cov}(x_i\gamma, \bar{x}\gamma) = c_i\mathrm{Var}(\bar{x}\gamma)$$

$$\iff c_i = \frac{\mathrm{Cov}(\beta_i, \beta_{ATE})}{\mathrm{Var}(\beta_{ATE})}$$

The random variable in this context is γ (the coefficients on the x's) in this case $\mathrm{Var}(\beta_{ATE})$ is a measure of how uncertain one was about what β_{ATE} would be before doing the trial.

c_i is the equation for a regression of β_i on β_{ATE}. In other words, we take a bunch of patients with characteristics x_i and we keep redrawing the gammas from our prior distribution. Then we ask how correlated β_i and β_{ATE} are. If they are more correlated (as they would be for patients where the x_i are closer to $\bar{x}$) we update more.

To compute c_i, we just need to know x_i, $\bar{x}$, and the distribution of γ.

Suppose we want to design the trial to minimize:

$$\min \mathrm{E}[(\beta_i - \beta_{i,post})^2]$$

2 Results

We start by considering two subpopulations: men and women. In this simple scenario, x_i is an indicator for whether individual i is a man. By symmetry, $x_i = 0$ implies individual i is a woman. In addition, we let p be the proportion of men in the population and $\bar{x}$ to be the fraction of men in the trial.

Below we will consider the results for four different cases and observe how the results for the optimal trial composition differ. The first two cases assume the covariance in treatment effects for men and women is zero. In cases 3 and 4 we generalize the first two cases by allowing for arbitrary covariance between the treatment effects for men and women.

In all the cases, we say a solution is optimal if it minimizes the mean squared error of $(\beta_{i,post} - \beta_i)$

Case 1

Our first setup is

$$\beta_i = (1 - x)\gamma_0 + x\gamma_1$$

$$\beta_{ATE} = (1 - \bar{x})\gamma_0 + \bar{x}\gamma_1$$

When looking at the derivations we observe that the optimal representation in the clinical trial depends both on the proportion of men and women in the population and their respective variances in treatment effects. Interestingly, when the proportion of men in the population is greater than the ratio of the variance in treatment effects for women to the population, (e.g. $p > \frac{\text{Var}(\gamma_0)}{\text{Var}(\gamma_0)+\text{Var}(\gamma_1)}$), it is optimal to choose only men to be in our trial. Conversely, if $p < \frac{\text{Var}(\gamma_0)}{\text{Var}(\gamma_0)+\text{Var}(\gamma_1)}$ we choose only women. The final case is if $p = \frac{\text{Var}(\gamma_0)}{\text{Var}(\gamma_0)+\text{Var}(\gamma_1)}$. In this scenario it doesn't matter what proportions we include in our trial, the MSE will be the same for all $\bar{x}$.

If we impose that $\text{Var}(\gamma_0) = \text{Var}(\gamma_1) = 1$ then we get that when $p > \frac{1}{2}$ we only want men in the clinical trial. If $p < \frac{1}{2}$, we only want women in the clinical trial. And if $p = \frac{1}{2}$, then any choice of composition of men and women in the clinical trial will lead to the same MSE.

Case 2

For Case 2, we have

$$\beta_i = \gamma_0 + x\gamma_1$$

$$\beta_{ATE} = \gamma_0 + \bar{x}\gamma_1$$

In this case, we observe the same general rule that the optimal proportion in the clinical trial depends on the proportion of men and women in the population and their respective variances in outcomes. However, because of the structure of β_i, in this setup men always have a higher variance in outcomes.

This asymmetric variance between men and women leads to the scenario where there is always a preference for a greater fraction of the clinical trial to be men than in the population. For example, if the true population contains 50% men ($p = \frac{1}{2}$), then we would want $\approx 61.8\%$ of the trial to be men in order to minimize the MSE.

To further investigate the effect of variance in this case we can observe that as $\text{Var}(\gamma_1) \to \infty$ holding $\text{Var}(\gamma_0)$ fixed we always choose only men. And as the variance for women $\text{Var}(\gamma_0) \to \infty$ holding $\text{Var}(\gamma_1)$ fixed, we choose $\bar{x} = p$. In other words, the optimal clinical trial composition is simply the true population proportion of men and women.

Case 3

In this case, we generalize case 2 to allow for arbitrary covariance between γ_0 and γ_1

Recall that our setup in case 2 is

$$\beta_i = \gamma_0 + x\gamma_1$$

$$\beta_{ATE} = \gamma_0 + \bar{x}\gamma_1$$

First, we can see in 3 that when we let $\text{Cov}(\gamma_0, \gamma_1) = 0$ we recover case 2.

Furthermore, our intuition from the previous results suggest that the representation depends on the respective variances in outcomes of each group.

We will thus consider a set of options.

Given the setup, we can observe that

$$\text{Var}(\beta_{women}) = \text{Var}(\gamma_0)$$

$$\text{Var}(\beta_{men}) = \text{Var}(\gamma_0) + 2\text{Cov}(\gamma_0, \gamma_1) + \text{Var}(\gamma_1)$$

We will impose that $\text{Var}(\gamma_0) = \text{Var}(\gamma_1) = 1$ and consider different covariances to make men or women have greater variance in outcomes.

From the variances for outcomes above, we can see that for men and women to have equal variance we must impose that $\text{Cov}(\gamma_0, \gamma_1) = -\frac{1}{2}$.

We investigate different covariance options in figure 1. In the figure we observe that if men and women have the same variance in outcomes, then you over represent whichever group is in the majority. As the variance of one group increases relative to another, you choose to over represent that group even more. And finally, as seen in the last graph of the figure, we can replicate the results from case 1 where it is optimal to only include the majority group in our trial by making the outcomes for men and women inversely correlated while having the same variance.

Case 4

In this case, we generalize case 1 to allow for arbitrary correlation between γ_0 and γ_1. Recall that in case 1 we had

$$\beta_i = (1 - x)\gamma_0 + x\gamma_1$$

$$\beta_{ATE} = (1 - \bar{x})\gamma_0 + \bar{x}\gamma_1$$

First, we can see in 3 that when we let $\text{Cov}(\gamma_0, \gamma_1) = 0$ we recover case 1.

As before, we will consider a range of options changing the respective variances in outcomes for men and women.

Given the setup, we can observe that

$$\text{Var}(\beta_{women}) = \text{Var}(\gamma_0)$$

$$\text{Var}(\beta_{men}) = \text{Var}(\gamma_1)$$

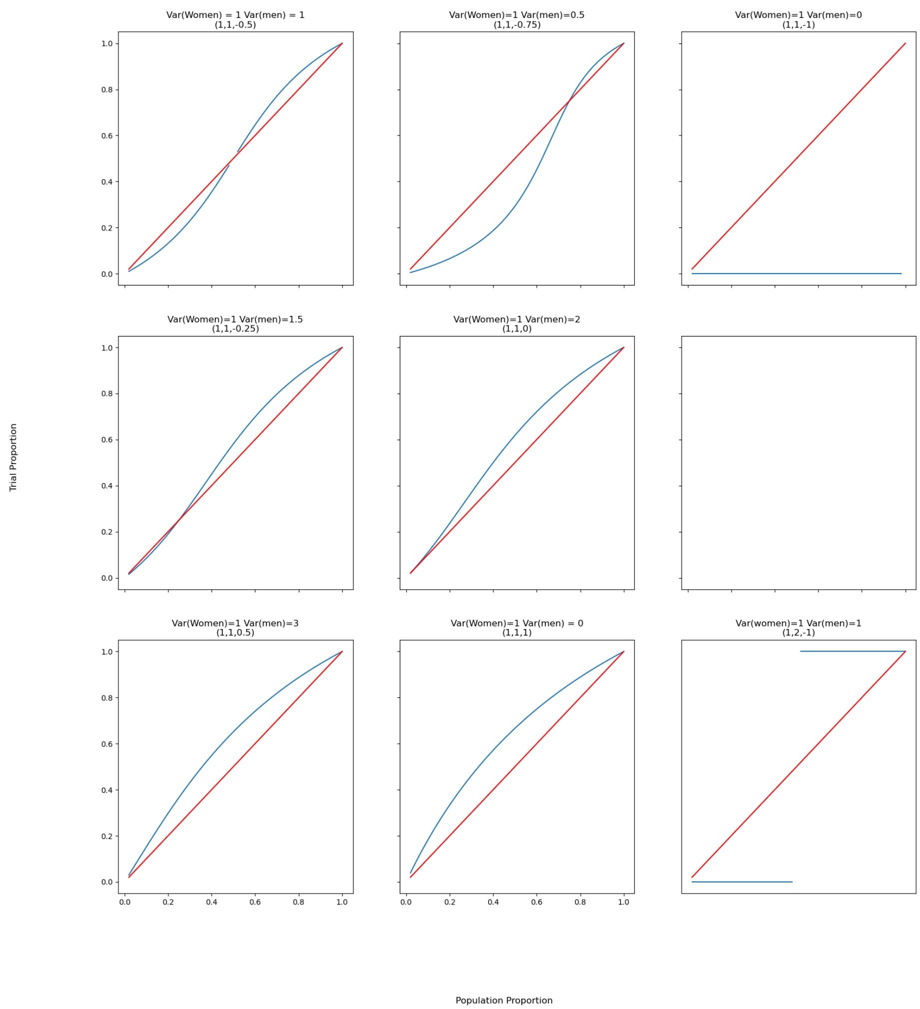

Figure 1: The figures contain different options for $\mathrm{Cov}(\gamma_0, \gamma_1)$ and for all but the last graph, $\mathrm{Var}(\gamma_0) = \mathrm{Var}(\gamma_1) = 1$

In this setup, the covariance structure doesn't tell us which group has more variance. However it does tell us how much the outcomes for men tells us about the distribution of outcomes for women and vice versa. With this interpretation, we can see that as $\mathrm{Cov}(\gamma_0, \gamma_1) \to 1$ the optimal study proportions $\bar{x} \to p$. In other words, the optimal proportion of men in the trial is the proportion of men in the population.

As $\mathrm{Cov}(\gamma_0, \gamma_1) \to 0$, we recover the results from case 1 where the equal variances tell us that because no group has greater variance, whichever group is more common in the true population should be maximized in the study. e.g. if men are more than half of the population ($p > \frac{1}{2}$) then we should only include men in the clinical trial ($\bar{x} = 1$) and vice versa.

Finally, if $\mathrm{Cov}(\gamma_0, \gamma_1) \in (-1, 0)$ then we get edge cases where if men are the majority ($p > \frac{1}{2}$), then MSE is minimized at $\bar{x} = 1$ and vice versa. If $p = \frac{1}{2}$ then it is optimal to choose either all women or all men to be in the clinical trial. Choosing only to include men or women produces the same minimizing MSE. And as $\mathrm{Cov}(\gamma_0, \gamma_1) \to -1$, MSE converges point wise to 0 with a discontinuity at $\frac{1}{2}$. This means that it does not matter what the proportions of men and women are in the true population, any composition of men and women in the clinical trial will produce the same minimizing MSE.

We can see these results in figure 2

3 Appendix

Derivations

1. symmetric case

For this case, let $\beta_i = (1-x)\gamma_0 + x\gamma_1$ and define β_{ATE} likewise.

Recall we want to minimize

$$\min \mathrm{E}_x \left(\mathrm{E}_{\gamma_{0,1}}[(\beta_i - \beta_{i,post})^2] \right)$$

One way we can rewrite these equations is as the effect of women vs men. let the squared error (SE) be

$$\sqrt{SE} = \beta_i - \beta_{i,post} = \underbrace{[(1-x)\gamma_0 + x\gamma_1]}_{\beta_i} - \underbrace{[((\vec{x} - c\vec{\bar{x}})\mathrm{E}(\gamma) + c\beta_{ATE})]}_{\beta_{i,post}}$$

$$= [(1-x)\gamma_0 + x\gamma_1] - [((1-x) - c(1-\bar{x}))\mathrm{E}(\gamma_0) + (x - c\bar{x})\mathrm{E}(\gamma_1) + c((1-\bar{x})\gamma_0 + \bar{x}\gamma_1)]$$

Because this is a function of x, We can then describe
$\beta_i^{men} = \beta_i(x = 1)$ and similarly for other terms to get

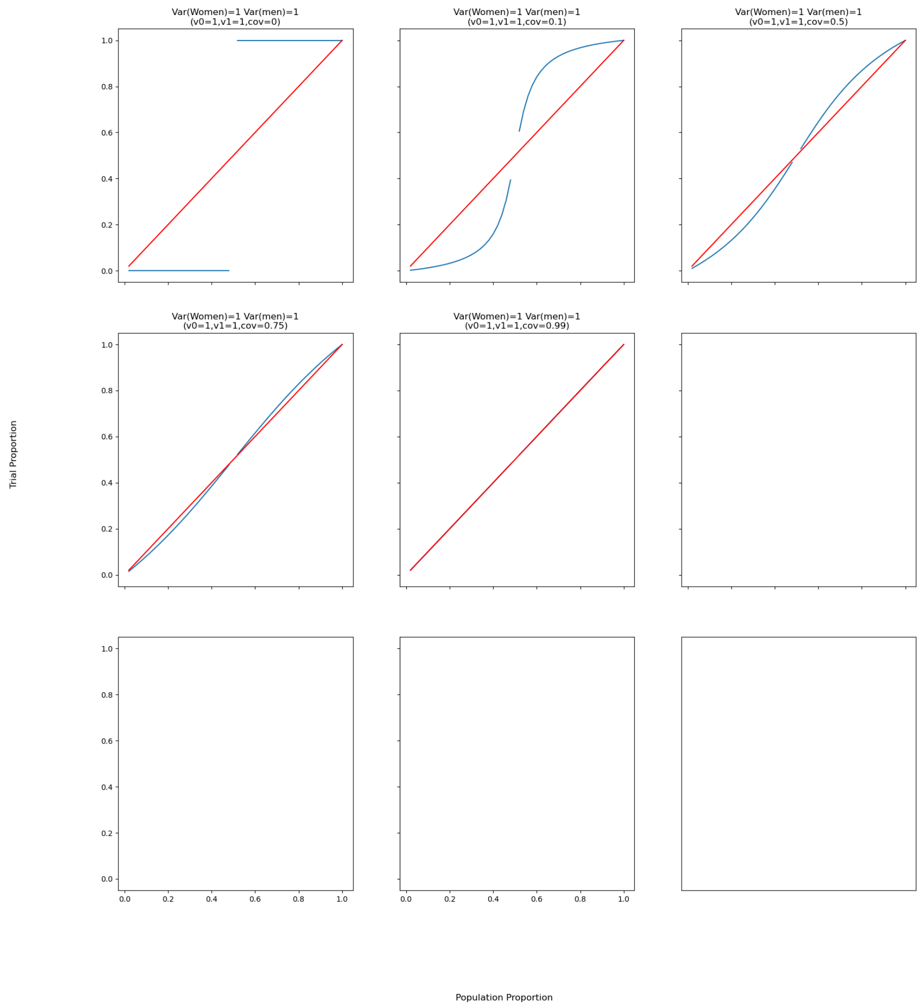

Figure 2: Here we show the optimal trial proportion of men vs the population proportion (on the y and x axis respectively) for equal variances of 1.

So

$$\beta_i^{men} = \gamma_1$$

$$\beta_i^{post,men} = -c_{men}(1-\bar{x})\bar{\gamma}_0 + (1 - c_{men}\bar{x})\bar{\gamma}_1 + c_{men}[(1-\bar{x})\gamma_0 + \bar{x}\gamma_1]$$

$$\beta_i^{women} = \gamma_0$$

$$\beta_i^{post,women} = [1 - c_{wom}(1-\bar{x})]\mathrm{E}(\gamma_0) - c_{wom}\bar{x}\mathrm{E}(\gamma_1) + c_{wom}(1-\bar{x})\gamma_0 + c_{wom}\bar{x}\gamma_1$$

this can be broken down into

$$\sqrt{SE} = (\beta_i^{men} - \beta_i^{post,men}) + (\beta_i^{wom} - \beta_i^{post,wom})$$

let

$$\begin{aligned}
W^{men} &= (\beta_i^{men} - \beta_i^{post,men}) \\
&= (1 - \bar{x}c_{men})(\gamma_1 - \bar{\gamma}_1) - c_{men}(1 - \bar{x})(\gamma_0 - \bar{\gamma}_0) \\
&= (c_{wom}(1 - \bar{x}))(\gamma_1 - \bar{\gamma}_1) - c_{men}(1 - \bar{x})(\gamma_0 - \bar{\gamma}_0) \\
W^{wom} &= (\beta_i^{wom} - \beta_i^{post,wom}) \\
&= [1 - c_{wom}(1 - \bar{x})](\gamma_0 - \bar{\gamma}_0) - c_{wom}\bar{x}(\gamma_1 - \bar{\gamma}_1) \\
&= [c_{men}\bar{x}](\gamma_0 - \bar{\gamma}_0) - c_{wom}\bar{x}(\gamma_1 - \bar{\gamma}_1)
\end{aligned}$$

and recall that

$$c_i = \frac{(1-x)(1-\bar{x})\mathrm{Var}(\gamma_0) + x\bar{x}\mathrm{Var}(\gamma_1)}{(1-\bar{x})^2\mathrm{Var}(\gamma_0) + \bar{x}^2\mathrm{Var}(\gamma_1)}$$

so if you are a man, then

$$c_{men} = \frac{\bar{x}\mathrm{Var}(\gamma_1)}{(1-\bar{x})^2\mathrm{Var}(\gamma_0) + \bar{x}^2\mathrm{Var}(\gamma_1)}$$

and likewise if you are a woman, then

$$c_{woman} = \frac{(1-\bar{x})\mathrm{Var}(\gamma_0)}{(1-\bar{x})^2\mathrm{Var}(\gamma_0) + \bar{x}^2\mathrm{Var}(\gamma_1)}$$

And, allowing for arbitrary variances, we can say that

$$W_{men}^2 = \left(\frac{(1-\bar{x})^4}{((1-\bar{x})^2\text{Var}(\gamma_0)+\bar{x}^2\text{Var}(\gamma_1))^2}\right)\text{Var}(\gamma_0)^2\text{Var}(\gamma_1)$$

$$- \frac{2(1-\bar{x})^2[\bar{x}(1-\bar{x})]}{((1-\bar{x})^2\text{Var}(\gamma_0)+\bar{x}^2\text{Var}(\gamma_1))^2}\text{Var}(\gamma_1)\text{Var}(\gamma_0)\text{Cov}(\gamma_0,\gamma_1)$$

$$+ \frac{\bar{x}^2(1-\bar{x})^2}{((1-\bar{x})^2\text{Var}(\gamma_0)+\bar{x}^2\text{Var}(\gamma_1))^2}\text{Var}(\gamma_1)^2\text{Var}(\gamma_0)$$

$$= \text{Var}(\gamma_0)\text{Var}(\gamma_1)\bigg(\frac{(1-\bar{x})^4}{((1-\bar{x})^2\text{Var}(\gamma_0)+\bar{x}^2\text{Var}(\gamma_1))^2}\text{Var}(\gamma_0)$$

$$- \frac{2(1-\bar{x})^2[\bar{x}(1-\bar{x})]}{((1-\bar{x})^2\text{Var}(\gamma_0)+\bar{x}^2\text{Var}(\gamma_1))^2}\text{Cov}(\gamma_0,\gamma_1)$$

$$+ \frac{\bar{x}^2(1-\bar{x})^2}{((1-\bar{x})^2\text{Var}(\gamma_0)+\bar{x}^2\text{Var}(\gamma_1))^2}\text{Var}(\gamma_1)\bigg)$$

$$W_{women}^2 = \left(\frac{(\bar{x})^4}{((1-\bar{x})^2\text{Var}(\gamma_0)+\bar{x}^2\text{Var}(\gamma_1))^2}\right)\text{Var}(\gamma_1)^2\text{Var}(\gamma_0)$$

$$- \frac{2\bar{x}^2[\bar{x}(1-\bar{x})]}{((1-\bar{x})^2\text{Var}(\gamma_0)+\bar{x}^2\text{Var}(\gamma_1))^2}\text{Var}(\gamma_1)\text{Var}(\gamma_0)\text{Cov}(\gamma_0,\gamma_1)$$

$$+ \frac{\bar{x}^2(1-\bar{x})^2}{((1-\bar{x})^2\text{Var}(\gamma_0)+\bar{x}^2\text{Var}(\gamma_1))^2}\text{Var}(\gamma_0)^2\text{Var}(\gamma_1)$$

$$= \text{Var}(\gamma_0)\text{Var}(\gamma_1)\bigg(\frac{(\bar{x})^4}{((1-\bar{x})^2\text{Var}(\gamma_0)+\bar{x}^2\text{Var}(\gamma_1))^2}\text{Var}(\gamma_1)$$

$$- \frac{2\bar{x}^2[\bar{x}(1-\bar{x})]}{((1-\bar{x})^2\text{Var}(\gamma_0)+\bar{x}^2\text{Var}(\gamma_1))^2}\text{Cov}(\gamma_0,\gamma_1)$$

$$+ \frac{\bar{x}^2(1-\bar{x})^2}{((1-\bar{x})^2\text{Var}(\gamma_0)+\bar{x}^2\text{Var}(\gamma_1))^2}\text{Var}(\gamma_0)\bigg)$$

So letting $\text{Cov}(\gamma_0,\gamma_1) = 0$ we get

$$W_{men}^2 = \text{Var}(\gamma_0)\text{Var}(\gamma_1)\left(\frac{(1-\bar{x})^2}{(1-\bar{x})^2\text{Var}(\gamma_0)+\bar{x}^2\text{Var}(\gamma_1)}\right)$$

$$W_{women}^2 = \text{Var}(\gamma_0)\text{Var}(\gamma_1)\left(\frac{(\bar{x})^2}{(1-\bar{x})^2\text{Var}(\gamma_0)+\bar{x}^2\text{Var}(\gamma_1)}\right)$$

Now observe that if $\alpha = \text{Var}(\gamma_1)$ and $\beta = \text{Var}(\gamma_0)$ then

$$\alpha W_{women}^2 + \beta W_{men}^2 = \text{Var}(\gamma_0)\text{Var}(\gamma_1)\left[\frac{\text{Var}(\gamma_1)\bar{x}^2 + \text{Var}(\gamma_0)(1-\bar{x})^2}{(1-\bar{x})^2\text{Var}(\gamma_0)+\bar{x}^2\text{Var}(\gamma_1)}\right]$$

$$= \text{Var}(\gamma_0)\text{Var}(\gamma_1)$$

So we can write

$$\alpha W_{women}^2 = \text{Var}(\gamma_0)\text{Var}(\gamma_1) - \beta W_{men}^2$$

$$W_{women}^2 = \frac{\text{Var}(\gamma_0)\text{Var}(\gamma_1) - \beta W_{men}^2}{\alpha}$$

$$W_{women}^2 = \frac{\text{Var}(\gamma_0)\text{Var}(\gamma_1) - \text{Var}(\gamma_0) W_{men}^2}{\text{Var}(\gamma_1)}$$

And thus conclude

$$MSE = (1-p)W_{women}^2 + pW_{men}^2$$

$$MSE = pW_{men}^2 + (1-p)\left(\frac{\text{Var}(\gamma_0)\text{Var}(\gamma_1) - \text{Var}(\gamma_0) W_{men}^2}{\text{Var}(\gamma_1)}\right)$$

$$MSE = pW_{men}^2 + (1-p)\left(\frac{\text{Var}(\gamma_0)\text{Var}(\gamma_1) - \text{Var}(\gamma_0) W_{men}^2}{\text{Var}(\gamma_1)}\right)$$

$$= \frac{\left(p[\text{Var}(\gamma_0)+\text{Var}(\gamma_1)] - \text{Var}(\gamma_0)\right)W_{men}^2 + (1-p)\text{Var}(\gamma_0)\text{Var}(\gamma_1)}{\text{Var}(\gamma_1)}$$

Thus when $p > \frac{\text{Var}(\gamma_0)}{\text{Var}(\gamma_0)+\text{Var}(\gamma_1)}$ we can clearly see that MSE is minimized when W_{men}^2 is minimized (when $\bar{x} = 1$). And inversely when $p < \frac{\text{Var}(\gamma_0)}{\text{Var}(\gamma_0)+\text{Var}(\gamma_1)}$ MSE is minimized when W_{men}^2 is maximaximized (when $\bar{x}=0$). In other words, when the proportion of men is $p > \frac{\mathrm{Var}(\gamma_0)}{\mathrm{Var}(\gamma_0)+\mathrm{Var}(\gamma_1)}$ it is optimal to have only men ($\bar{x}$) in the trial, and vice versa. And when there are eqal number of men and women in the population, MSE does not depend on W^2_{men} and thus has equal error of $\frac{\mathrm{Var}(\gamma_0)}{\mathrm{Var}(\gamma_0)+\mathrm{Var}(\gamma_1)}$ for any $\bar{x}$.

We can also solve this using the first order conditions.

$$MSE = (1-p)W^2_{women} + pW^2_{men} = \mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1)\Big[\frac{(1-p)\bar{x}^2 + p(1-\bar{x})^2}{(1-\bar{x})^2\mathrm{Var}(\gamma_0) + \bar{x}^2\mathrm{Var}(\gamma_1)}\Big]$$

Which gives a derivative of

$$\frac{MSE}{d\bar{x}} = -\frac{2\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1)(1-\bar{x})\bar{x}\big[\mathrm{Var}(\gamma_0)(p-1) + \mathrm{Var}(\gamma_1)p\big]}{(\mathrm{Var}(\gamma_0)(\bar{x}-1)^2 + \mathrm{Var}(\gamma_1)\bar{x}^2)^2}$$

Here we can see that we still have roots $\bar{x}=0$ and $\bar{x}=1$ and all that changes is we get a variances-weighted edge-case whenever $(\mathrm{Var}(\gamma_0)+\mathrm{Var}(\gamma_1))p - \mathrm{Var}(\gamma_0) = 0$

And to get which is the minizing solution we can observe the SOC

$$\frac{MSE}{d\bar{x}^2} = -\frac{2\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1)[\mathrm{Var}(\gamma_0)(p-1)+\mathrm{Var}(\gamma_1)p][\mathrm{Var}(\gamma_0)(2x+1)(x-1)^2 + \mathrm{Var}(\gamma_1)x^2(2x-3)]}{[\mathrm{Var}(\gamma_0)(x-1)^2 + \mathrm{Var}(\gamma_1)x^2]^3}$$

From the second order condition, we can look at the two roots. We can see that the numerator reduces to two cases. When $p > \frac{\mathrm{Var}(\gamma_0)}{(\mathrm{Var}(\gamma_0)+\mathrm{Var}(\gamma_1))}$, then we can observe that the numerator is positive for the root $\bar{x}=1$. Conversely, the numerator is positive when $\bar{x}=0$ We can also see this outlined in the simulations below.

2. Men Have Higher Variance

$$\begin{aligned}
\beta_i &= \gamma_0 + x\gamma_1 \\
\beta_i^{post} &= \mathrm{E}(\gamma_0) + \mathrm{E}(\gamma_1) + c_i(\gamma_0 + \bar{x}\gamma_1 - (\mathrm{E}(\gamma_0) + \bar{x}\mathrm{E}(\gamma_1))) \\
\beta_i^{women} &= \gamma_0 \\
\beta_i^{men} &= \gamma_0 + \gamma_1 \\
\beta_i^{post,men} &= \mathrm{E}(\gamma_0) + \mathrm{E}(\gamma_1) + c_{men}(\gamma_0 + \bar{x}\gamma_1 - (\mathrm{E}(\gamma_0) + \bar{x}\mathrm{E}(\gamma_1))) \\
\beta_i^{post,women} &= \mathrm{E}(\gamma_0) + c_{women}(\gamma_0 + \bar{x}\gamma_1 - (\mathrm{E}(\gamma_0) + \bar{x}\mathrm{E}(\gamma_1)))
\end{aligned}$$

And we have

$$C_i = \frac{\mathrm{Var}(\gamma_0) + x\bar{x}\mathrm{Var}(\gamma_1)}{\mathrm{Var}(\gamma_0) + \bar{x}^2\mathrm{Var}\gamma_1)}$$

$$C_{women} = \frac{\mathrm{Var}(\gamma_0)}{\mathrm{Var}(\gamma_0) + \bar{x}^2\mathrm{Var}(\gamma_1)}$$

$$C_{men} = \frac{\mathrm{Var}(\gamma_0) + \bar{x}\mathrm{Var}(\gamma_1)}{\mathrm{Var}(\gamma_0) + \bar{x}^2\mathrm{Var}(\gamma_1)}$$

$$\begin{aligned}
W_{women} &= \beta_i^{women} - \beta_i^{post,women} \\
&= (\gamma_0 - \mathrm{E}(\gamma_0)) - C_{men}\big[(\gamma_0 - \mathrm{E}(\gamma_0)) + \bar{x}(\gamma_1 - \mathrm{E}(\gamma_1))\big] \\
&= (1 - c_{women})(\gamma_0 - \mathrm{E}(\gamma_0)) - C_{women}\bar{x}(\gamma_1 - \mathrm{E}(\gamma_1)) \\
W_{men} &= \beta_i^{men} - \beta_i^{post,men} \\
&= (\gamma_0 - \mathrm{E}(\gamma_0)) + (\gamma_1 - \mathrm{E}(\gamma_1) - C_{men}\big[(\gamma_0 - \mathrm{E}(\gamma_0)) + \bar{x}(\gamma_1 - \mathrm{E}(\gamma_1))\big] \\
&= (1 - c_{men})(\gamma_0 - \mathrm{E}(\gamma_0)) + (1 - C_{men}\bar{x})(\gamma_1 - \mathrm{E}(\gamma_1))
\end{aligned}$$

Squaring

$$W_{women}^2 = (1 - c_{wom})^2\mathrm{Var}(\gamma_0) - 2(1 - c_{wom})c_{wom}\bar{x}\mathrm{Cov}(\gamma_0, \gamma_1) + c_{wom}^2\bar{x}\mathrm{Var}(\gamma_1)$$

$$W_{men}^2 = (1 - c_{men})^2\mathrm{Var}(\gamma_0) + 2(1 - c_{men})(1 - c_{men}\bar{x})\mathrm{Cov}(\gamma_0, \gamma_1) + (1 - c_{men}\bar{x})^2\mathrm{Var}(\gamma_1)$$

And assuming $\mathrm{Cov}(\gamma_0, \gamma_1) = 0$ we get

$$
\begin{aligned}
W^2_{women} &= (1-c_{wom})^2\mathrm{Var}(\gamma_0) + c^2_{wom}\bar{x}^2\mathrm{Var}(\gamma_1) \\
&= \frac{\bar{x}^4\mathrm{Var}(\gamma_1)^2}{(\mathrm{Var}(\gamma_0)+\bar{x}^2\mathrm{Var}(\gamma_1))^2}\mathrm{Var}(\gamma_0) + \frac{\bar{x}^2\mathrm{Var}(\gamma_0)^2}{(\mathrm{Var}(\gamma_0)+\bar{x}^2\mathrm{Var}(\gamma_1))^2}\mathrm{Var}(\gamma_1) \\
&= \bar{x}^2\mathrm{Var}(\gamma_1)\mathrm{Var}(\gamma_0)\frac{\bar{x}^2\mathrm{Var}(\gamma_1)+\mathrm{Var}(\gamma_0)}{(\mathrm{Var}(\gamma_0)+\bar{x}^2\mathrm{Var}(\gamma_1))^2} \\
&= \frac{\bar{x}^2\mathrm{Var}(\gamma_1)\mathrm{Var}(\gamma_0)}{\mathrm{Var}(\gamma_0)+\bar{x}^2\mathrm{Var}(\gamma_1)} \\
W^2_{men} &= (1-c_{men})^2\mathrm{Var}(\gamma_0) + (1-c_{men}\bar{x})^2\mathrm{Var}(\gamma_1) \\
&= \frac{\mathrm{Var}(\gamma_1)^2[\bar{x}^2-\bar{x}]^2}{(\mathrm{Var}(\gamma_0)+\bar{x}^2\mathrm{Var}(\gamma_1))^2}\mathrm{Var}(\gamma_0) + \frac{\mathrm{Var}(\gamma_0)^2[1-\bar{x}]^2}{(\mathrm{Var}(\gamma_0)+\bar{x}^2\mathrm{Var}(\gamma_1))^2}\mathrm{Var}(\gamma_1) \\
&= \frac{\mathrm{Var}(\gamma_1)^2\bar{x}^2[\bar{x}-1]^2}{(\mathrm{Var}(\gamma_0)+\bar{x}^2\mathrm{Var}(\gamma_1))^2}\mathrm{Var}(\gamma_0) + \frac{\mathrm{Var}(\gamma_0)^2[\bar{x}-1]^2}{(\mathrm{Var}(\gamma_0)+\bar{x}^2\mathrm{Var}(\gamma_1))^2}\mathrm{Var}(\gamma_1) \\
&= \mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1)[\bar{x}-1]^2\frac{\mathrm{Var}(\gamma_1)\bar{x}^2+\mathrm{Var}(\gamma_0)}{(\mathrm{Var}(\gamma_0)+\bar{x}^2\mathrm{Var}(\gamma_1))^2} \\
&= \frac{\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1)[\bar{x}-1]^2}{\mathrm{Var}(\gamma_0)+\bar{x}^2\mathrm{Var}(\gamma_1)}
\end{aligned}
$$

So we get MSE of

$$MSE = \mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1)\frac{(1-p)\bar{x}^2+p[\bar{x}-1]^2}{\mathrm{Var}(\gamma_0)+\bar{x}^2\mathrm{Var}(\gamma_1)}$$

Taking first order conditions we get

$$
\begin{aligned}
&\frac{d}{dx}\frac{\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1)[(1-p)x^2+p(x-1)^2]}{\mathrm{Var}(\gamma_0)+\mathrm{Var}(\gamma_1)x^2} \\
&\quad = 2\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1)\frac{\mathrm{Var}(\gamma_0)(x-p)+\mathrm{Var}(\gamma_1)p(x-1)x}{(\mathrm{Var}(\gamma_0)+\mathrm{Var}(\gamma_1)x^2)^2} = 0 \\
&\quad \iff \mathrm{Var}(\gamma_0)(x-p)+\mathrm{Var}(\gamma_1)p(x-1)x = 0 \\
&\quad \iff x^2\mathrm{Var}(\gamma_1)p + x(\mathrm{Var}(\gamma_0)-\mathrm{Var}(\gamma_1)p) - \mathrm{Var}(\gamma_0)p = 0
\end{aligned}
$$

For this we can use the quadratic formula with

1. $a = \text{Var}(\gamma_1)p$
2. $b = \text{Var}(\gamma_0) - \text{Var}(\gamma_1)p$
3. $c = -\text{Var}(\gamma_0)p$

giving us

$$\begin{aligned} x &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \\ &= \frac{-(\text{Var}(\gamma_0) - \text{Var}(\gamma_1)p) \pm \sqrt{(\text{Var}(\gamma_0) - \text{Var}(\gamma_1)p)^2 + 4\text{Var}(\gamma_0)\text{Var}(\gamma_1)p^2}}{2\text{Var}(\gamma_1)p} \\ &= \frac{\text{Var}(\gamma_1)p - \text{Var}(\gamma_0) \pm \sqrt{\text{Var}(\gamma_0)^2 - 2\text{Var}(\gamma_0)\text{Var}(\gamma_1)p + \text{Var}(\gamma_1)^2p^2 + 4\text{Var}(\gamma_0)\text{Var}(\gamma_1)p^2}}{2\text{Var}(\gamma_1)p} \\ &= \frac{\text{Var}(\gamma_1)p - \text{Var}(\gamma_0) \pm \sqrt{\text{Var}(\gamma_0)^2 + 2\text{Var}(\gamma_0)\text{Var}(\gamma_1)p(2p-1) + \text{Var}(\gamma_1)^2p^2}}{2\text{Var}(\gamma_1)p} \end{aligned}$$

We can see the errors and optimal solutions 4 and 5

3. Generalization of case 1 and 2

In case 1, we solved

$$\beta_i = (1-x)\gamma_0 + x\gamma_1 = \gamma_0 + x(\gamma_1 - \gamma_0)$$

Thus in effect, the constant and coefficient are correlated through γ_0. We can thus nest both cases by allowing for arbitrary correlation in case 1.

We will do this below

First for C

We have that

$$c = \frac{\text{Cov}(x\gamma, \bar{x}\gamma)}{\text{Var}(\bar{x}\gamma)}$$

so for individual i, c reduces to

$$c = \frac{\text{Var}(\gamma_0) + (x + \bar{x})\text{Cov}(\gamma_0, \gamma_1) + x\bar{x}\text{Var}(\gamma_1)}{\text{Var}(\gamma_0) + 2\bar{x}\text{Cov}(\gamma_0, \gamma_1) + \bar{x}^2\text{Var}(\gamma_1)}$$

$$C_{women} = \frac{\text{Var}(\gamma_0) + \bar{x}\text{Cov}(\gamma_0, \gamma_1)}{\text{Var}(\gamma_0) + 2\bar{x}\text{Cov}(\gamma_0, \gamma_1) + \bar{x}^2\text{Var}(\gamma_1)}$$

$$C_{men} = \frac{\text{Var}(\gamma_0) + (1 + \bar{x})\text{Cov}(\gamma_0, \gamma_1) + \bar{x}\text{Var}(\gamma_1)}{\text{Var}(\gamma_0) + 2\bar{x}\text{Cov}(\gamma_0, \gamma_1) + \bar{x}^2\text{Var}(\gamma_1)}$$

We still have

$$\begin{aligned}
\beta_i &= \gamma_0 + x\gamma_1 \\
\beta_i^{post} &= \text{E}(\gamma_0) + x\text{E}(\gamma_1) + c_i(\gamma_0 + \bar{x}\gamma_1 - (\text{E}(\gamma_0) + \bar{x}\text{E}(\gamma_1))) \\
\beta_i^{women} &= \gamma_0 \\
\beta_i^{men} &= \gamma_0 + \gamma_1 \\
\beta_i^{post,men} &= \text{E}(\gamma_0) + \text{E}(\gamma_1) + c_{men}(\gamma_0 + \bar{x}\gamma_1 - (\text{E}(\gamma_0) + \bar{x}\text{E}(\gamma_1))) \\
\beta_i^{post,women} &= \text{E}(\gamma_0) + c_{women}(\gamma_0 + \bar{x}\gamma_1 - (\text{E}(\gamma_0) + \bar{x}\text{E}(\gamma_1)))
\end{aligned}$$

and

$$\begin{aligned}
W_{women} &= \beta_i^{women} - \beta_i^{post,women} \\
&= (\gamma_0 - \text{E}(\gamma_0)) - C_{men}\big[(\gamma_0 - \text{E}(\gamma_0)) + \bar{x}(\gamma_1 - \text{E}(\gamma_1))\big] \\
&= (1 - c_{women})(\gamma_0 - \text{E}(\gamma_0)) - C_{women}\bar{x}(\gamma_1 - \text{E}(\gamma_1)) \\
W_{men} &= \beta_i^{men} - \beta_i^{post,men} \\
&= (\gamma_0 - \text{E}(\gamma_0)) + (\gamma_1 - \text{E}(\gamma_1) - C_{men}\big[(\gamma_0 - \text{E}(\gamma_0)) + \bar{x}(\gamma_1 - \text{E}(\gamma_1))\big] \\
&= (1 - c_{men})(\gamma_0 - \text{E}(\gamma_0)) + (1 - C_{men}\bar{x})(\gamma_1 - \text{E}(\gamma_1))
\end{aligned}$$

Squaring

$$W_{women}^2 = (1-c_{wom})^2\mathrm{Var}(\gamma_0) - 2(1-c_{wom})c_{wom}\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1) + c_{wom}^2\bar{x}^2\mathrm{Var}(\gamma_1)$$

$$= \frac{\big(\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)+\bar{x}^2\mathrm{Var}(\gamma_1)\big)^2}{\big(\mathrm{Var}(\gamma_0)+2\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)+\bar{x}^2\mathrm{Var}(\gamma_1)\big)^2}\mathrm{Var}(\gamma_0)-$$

$$2\left(\frac{[\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)+\bar{x}^2\mathrm{Var}(\gamma_1)][\mathrm{Var}(\gamma_0)+\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)]}{\big(\mathrm{Var}(\gamma_0)+2\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)+\bar{x}^2\mathrm{Var}(\gamma_1)\big)^2}\right)\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)+$$

$$\frac{\big(\mathrm{Var}(\gamma_0)+\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)\big)^2}{\big(\mathrm{Var}(\gamma_0)+2\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)+\bar{x}^2\mathrm{Var}(\gamma_1)\big)^2}\bar{x}^2\mathrm{Var}(\gamma_1)$$

$$= \frac{\bar{x}^2\mathrm{Cov}(\gamma_0,\gamma_1)^2+2\bar{x}^3\mathrm{Cov}(\gamma_0,\gamma_1)\mathrm{Var}(\gamma_1)+\bar{x}^4\mathrm{Var}(\gamma_1)^2}{\big(\mathrm{Var}(\gamma_0)+2\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)+\bar{x}^2\mathrm{Var}(\gamma_1)\big)^2}\mathrm{Var}(\gamma_0)$$

$$-2\left(\frac{\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)\mathrm{Var}(\gamma_0)+\bar{x}^2\mathrm{Cov}(\gamma_0,\gamma_1)^2+\bar{x}^2\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1)+\bar{x}^3\mathrm{Var}(\gamma_1)\mathrm{Cov}(\gamma_0,\gamma_1)}{\big(\mathrm{Var}(\gamma_0)+2\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)+\bar{x}^2\mathrm{Var}(\gamma_1)\big)^2}\right)\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1$$

$$+\frac{\mathrm{Var}(\gamma_0)^2+2\bar{x}\mathrm{Var}(\gamma_0)\mathrm{Cov}(\gamma_0,\gamma_1)+\bar{x}^2\mathrm{Cov}(\gamma_0,\gamma_1)^2}{\big(\mathrm{Var}(\gamma_0)+2\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)+\bar{x}^2\mathrm{Var}(\gamma_1)\big)^2}\bar{x}^2\mathrm{Var}(\gamma_1)$$

$$= \frac{\bar{x}^2\big[\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1)-\mathrm{Cov}(\gamma_0,\gamma_1)^2\big]}{\mathrm{Var}(\gamma_0)+\mathrm{Var}(\gamma_1)\bar{x}^2+2\mathrm{Cov}(\gamma_0,\gamma_1)\bar{x}}$$

$$W_{men}^2 = (1-c_{men})^2\mathrm{Var}(\gamma_0)+2(1-c_{men})(1-c_{men}\bar{x})\mathrm{Cov}(\gamma_0,\gamma_1)+(1-c_{men}\bar{x})^2\mathrm{Var}(\gamma_1)$$

$$= \frac{\big((\bar{x}-1)\mathrm{Cov}(\gamma_0,\gamma_1)+(\bar{x}^2-\bar{x})\mathrm{Var}(\gamma_1)\big)^2}{\big(\mathrm{Var}(\gamma_0)+2\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)+\bar{x}^2\mathrm{Var}(\gamma_1)\big)^2}\mathrm{Var}(\gamma_0)$$

$$+2\frac{[(\bar{x}-1)\mathrm{Cov}(\gamma_0,\gamma_1)+(\bar{x}^2-\bar{x})\mathrm{Var}(\gamma_1)][(1-\bar{x})\mathrm{Var}(\gamma_0)+(\bar{x}-\bar{x}^2)\mathrm{Cov}(\gamma_0,\gamma_1)]}{\big(\mathrm{Var}(\gamma_0)+2\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)+\bar{x}^2\mathrm{Var}(\gamma_1)\big)^2}\mathrm{Cov}(\gamma_0,\gamma_1)$$

$$+\frac{[(1-\bar{x})\mathrm{Var}(\gamma_0)+(\bar{x}-\bar{x}^2)\mathrm{Cov}(\gamma_0,\gamma_1)]^2}{\big(\mathrm{Var}(\gamma_0)+2\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)+\bar{x}^2\mathrm{Var}(\gamma_1)\big)^2}\mathrm{Var}(\gamma_1)$$

$$= \frac{(\bar{x}-1)^2\big[\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1)-\mathrm{Cov}(\gamma_0,\gamma_1)^2\big]}{\mathrm{Var}(\gamma_0)+\mathrm{Var}(\gamma_1)\bar{x}^2+2\mathrm{Cov}(\gamma_0,\gamma_1)\bar{x}}$$

We can see then that if $\mathrm{Cov}(\gamma_0,\gamma_1)=0$ we recover our result from version 2.
And we get a mean square error of

$$MSE = \big(\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1)-\mathrm{Cov}(\gamma_0,\gamma_1)^2\big)\frac{(1-p)\bar{x}^2+p(\bar{x}-1)^2}{\mathrm{Var}(\gamma_0)+\mathrm{Var}(\gamma_1)\bar{x}^2+2\mathrm{Cov}(\gamma_0,\gamma_1)\bar{x}}$$

$$\frac{d}{dx}\big(\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1) - \mathrm{Cov}(\gamma_0,\gamma_1)^2\big)\frac{(1-p)\bar{x}^2 + p(\bar{x}-1)^2}{\mathrm{Var}(\gamma_0) + \mathrm{Var}(\gamma_1)\bar{x}^2 + 2\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)}$$

$$= 2\big(\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1) - \mathrm{Cov}(\gamma_0,\gamma_1)^2\big)\frac{(\mathrm{Var}(\gamma_0)(\bar{x}-p) + \mathrm{Var}(\gamma_1)p(\bar{x}-1)\bar{x} + \mathrm{Cov}(\gamma_0,\gamma_1)(\bar{x}^2-p))}{(\mathrm{Var}(\gamma_0) + \bar{x}(\mathrm{Var}(\gamma_1)\bar{x} + 2\mathrm{Cov}(\gamma_0,\gamma_1)))^2}$$

Thus we can use the quadratic formula to determine the optimal trial proportions.

$$\bar{x} = \frac{\mathrm{Var}(\gamma_1)p - \mathrm{Var}(\gamma_0) \pm \sqrt{4p\big[\mathrm{Var}(\gamma_0) + \mathrm{Cov}(\gamma_0,\gamma_1)\big]\big[\mathrm{Var}(\gamma_1)p + \mathrm{Cov}(\gamma_0,\gamma_1)\big] + \big(\mathrm{Var}(\gamma_0) - \mathrm{Var}(\gamma_1)p\big)^2}}{2(\mathrm{Var}(\gamma_1)p + \mathrm{Cov}(\gamma_0,\gamma_1))}$$

and clearly $\mathrm{Var}(\gamma_1)p + \mathrm{Cov}(\gamma_0,\gamma_1) \neq 0$

We can thus see that we recover our original case of choosing between all men and all women whenever

$$\mathrm{Cov}(\gamma_0,\gamma_1) = -\mathrm{Var}(\gamma_0)$$

And when this holds, when $p = \frac{\mathrm{Var}(\gamma_0)}{\mathrm{Var}(\gamma_1)}$, we again get our edge-case where the the error is independent of the trial proportions and any $\bar{x}$ produces an equal error

Now we can check whether our setup in version 1 satisfies this.

Recall we have

$$\beta_i = \gamma_0 + x(\gamma_1 - \gamma_0)$$

Let's call $\gamma_2 = \gamma_1 - \gamma_0$

First observe that

$$\mathrm{Var}(\gamma_2) = \mathrm{Var}(\gamma_1 - \gamma_0) = \mathrm{Var}(\gamma_0) + \mathrm{Var}(\gamma_1) - 2\mathrm{Cov}(\gamma_0,\gamma_1) \quad \text{assuming both are 1}$$

$$= 1 + 1 - 0 = 2$$

and

$$\mathrm{Cov}(\gamma_0,\gamma_2) = \mathrm{Cov}(\gamma_0,\gamma_1-\gamma_0) = \mathrm{Cov}(\gamma_0,\gamma_1) - \mathrm{Cov}(\gamma_0,\gamma_0)$$

$$= 0 - 1 = -1$$

And so we recover that

$$\mathrm{Cov}(\gamma_0,\gamma_2) = -\mathrm{Var}(\gamma_0)$$

And more over that

$$p = \frac{\mathrm{Var}(\gamma_0)}{\mathrm{Var}(\gamma_2)} = \frac{1}{2}$$

And thus and MSE

$$MSE = \left(\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1) - \mathrm{Cov}(\gamma_0,\gamma_1)^2\right)\frac{(1-p)\bar{x}^2 + p(\bar{x}-1)^2}{\mathrm{Var}(\gamma_0) + \mathrm{Var}(\gamma_1)\bar{x}^2 + 2\mathrm{Cov}(\gamma_0,\gamma_1)\bar{x}}$$

$$= (2-1)\frac{\frac{1}{2}\bar{x}^2 + \frac{1}{2}(\bar{x}^2 - 2\bar{x} + 1)}{1 - 2\bar{x} + 2\bar{x}^2}$$

$$= \frac{\frac{1}{2}\bar{x}^2 + \frac{1}{2}(\bar{x}^2 - 2\bar{x} + 1)}{2(\frac{1}{2} - \bar{x} + \bar{x}^2)} = \frac{1}{2}$$

Which is always equal to $\frac{1}{2}$ and independent of $\bar{x}$. This is precisely what we get with the original solution to case 1.

4. Generalization of case 1 and 2

In case 1, we solved

$$\beta_i = (1-x)\gamma_0 + x\gamma_1 = \gamma_0 + x(\gamma_1 - \gamma_0)$$

Thus in effect, the constant and coefficient are correlated through γ_0. We can thus nest both cases by allowing for arbitrary correlation in case 1.

We will do this below

First for C

We have that

$$c = \frac{\mathrm{Cov}(x\gamma, \bar{x}\gamma)}{\mathrm{Var}(\bar{x}\gamma)}$$

so for individual i, c reduces to

$$c = \frac{(1-x)(1-\bar{x})\mathrm{Var}(\gamma_0) + x(1-\bar{x})\mathrm{Cov}(\gamma_0,\gamma_1) + (1-x)\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1) + x\bar{x}\mathrm{Var}(\gamma_1)}{(1-\bar{x})^2\mathrm{Var}(\gamma_0) + 2\bar{x}(1-\bar{x})\mathrm{Cov}(\gamma_0,\gamma_1) + \bar{x}^2\mathrm{Var}(\gamma_1)}$$

$$C_{women} = \frac{(1-\bar{x})\mathrm{Var}(\gamma_0) + \bar{x}\mathrm{Cov}(\gamma_0,\gamma_1)}{(1-\bar{x})^2\mathrm{Var}(\gamma_0) + 2(1-\bar{x})\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1) + \bar{x}^2\mathrm{Var}(\gamma_1)}$$

$$C_{men} = \frac{(1-\bar{x})\mathrm{Cov}(\gamma_0,\gamma_1) + \bar{x}\mathrm{Var}(\gamma_1)}{(1-\bar{x})^2\mathrm{Var}(\gamma_0) + 2(1-\bar{x})\bar{x}\mathrm{Cov}(\gamma_0,\gamma_1) + \bar{x}^2\mathrm{Var}(\gamma_1)}$$

We still have

$$
\begin{aligned}
\beta_i &= (1-x)\gamma_0 + x\gamma_1 \\
\beta_i^{post} &= (1-x)\mathrm{E}(\gamma_0) + x\mathrm{E}(\gamma_1) + c_i[(1-\bar{x})\gamma_0 + \bar{x}\gamma_1 - [(1-\bar{x})\mathrm{E}(\gamma_0) + \bar{x}\mathrm{E}(\gamma_1)]] \\
\beta_i^{women} &= \gamma_0 \\
\beta_i^{men} &= \gamma_1 \\
\beta_i^{post,men} &= \mathrm{E}(\gamma_1) + c_{men}((1-\bar{x})\gamma_0 + \bar{x}\gamma_1 - ((1-\bar{x})\mathrm{E}(\gamma_0) + \bar{x}\mathrm{E}(\gamma_1))) \\
\beta_i^{post,women} &= \mathrm{E}(\gamma_0) + c_{women}((1-\bar{x})\gamma_0 + \bar{x}\gamma_1 - ((1-\bar{x})\mathrm{E}(\gamma_0) + \bar{x}\mathrm{E}(\gamma_1)))
\end{aligned}
$$

and

$$
\begin{aligned}
W_{women} &= \beta_i^{women} - \beta_i^{post,women} \\
&= (\gamma_0 - \mathrm{E}(\gamma_0)) - C_{men}\big[(\gamma_0 - \mathrm{E}(\gamma_0)) + \bar{x}(\gamma_1 - \mathrm{E}(\gamma_1))\big] \\
&= (1 - c_{women}(1-\bar{x}))(\gamma_0 - \mathrm{E}(\gamma_0)) - C_{women}\bar{x}(\gamma_1 - \mathrm{E}(\gamma_1)) \\
&= c_{men}\bar{x}(\gamma_0 - \mathrm{E}(\gamma_0)) - C_{women}\bar{x}(\gamma_1 - \mathrm{E}(\gamma_1)) \\
W_{men} &= \beta_i^{men} - \beta_i^{post,men} \\
&= (\gamma_0 - \mathrm{E}(\gamma_0)) + (\gamma_1 - \mathrm{E}(\gamma_1) - C_{men}\big[(\gamma_0 - \mathrm{E}(\gamma_0)) + \bar{x}(\gamma_1 - \mathrm{E}(\gamma_1))\big] \\
&= (1 - c_{men}\bar{x})(\gamma_1 - \mathrm{E}(\gamma_1)) - C_{men}(1-\bar{x})(\gamma_0 - \mathrm{E}(\gamma_0)) \\
&= c_{women}(1-\bar{x})(\gamma_1 - \mathrm{E}(\gamma_1)) - C_{men}(1-\bar{x})(\gamma_0 - \mathrm{E}(\gamma_0))
\end{aligned}
$$

$$
\begin{aligned}
W_{women}^2 = {} & (C_{men}\bar{x})^2 \mathrm{Var}(\gamma_0) \\
& - 2\bar{x}^2 C_{women} C_{men} \mathrm{Cov}(\gamma_0, \gamma_1) \\
& + (\bar{x} C_{women})^2 \mathrm{Var}(\gamma_1) \\
= {} & \frac{(\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1) - \mathrm{Cov}(\gamma_0,\gamma_1)^2)\bar{x}^2}{\bar{x}^2(\mathrm{Var}(\gamma_0) + \mathrm{Var}(\gamma_1) - 2\mathrm{Cov}(\gamma_0,\gamma_1)) + 2\bar{x}(\mathrm{Cov}(\gamma_0,\gamma_1) - \mathrm{Var}(\gamma_0)) + \mathrm{Var}(\gamma_0)}
\end{aligned}
$$

$$
\begin{aligned}
W_{men}^2 &= (C_{women}(1-\bar{x}))^2 \mathrm{Var}(\gamma_1) \\
&\quad - 2(1-\bar{x})^2 C_{women} C_{men} \mathrm{Cov}(\gamma_0, \gamma_1) \\
&\quad + ((1-\bar{x})C_{men})^2 \mathrm{Var}(\gamma_0) \\
&= \frac{(\bar{x}-1)^2(\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1) - \mathrm{Cov}(\gamma_0,\gamma_1)^2)}{\bar{x}^2(\mathrm{Var}(\gamma_0) + \mathrm{Var}(\gamma_1) - 2\mathrm{Cov}(\gamma_0,\gamma_1)) + 2\bar{x}(\mathrm{Cov}(\gamma_0,\gamma_1) - \mathrm{Var}(\gamma_0)) + \mathrm{Var}(\gamma_0)}
\end{aligned}
$$

And so we get

$$
\begin{aligned}
MSE &= (1-p)W_{women}^2 + pW_{men}^2 \\
&= \frac{(\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1) - \mathrm{Cov}(\gamma_0,\gamma_1)^2)[(1-p)\bar{x}^2 + p(\bar{x}-1)^2]}{\bar{x}^2(\mathrm{Var}(\gamma_0) + \mathrm{Var}(\gamma_1) - 2\mathrm{Cov}(\gamma_0,\gamma_1)) + 2\bar{x}(\mathrm{Cov}(\gamma_0,\gamma_1) - \mathrm{Var}(\gamma_0)) + \mathrm{Var}(\gamma_0)}
\end{aligned}
$$

and derivative

$$
\frac{dMSE}{d\bar{x}} = \frac{2(\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1) - \mathrm{Cov}(\gamma_0,\gamma_1)^2)((\bar{x}-1)\bar{x}(\mathrm{Var}(\gamma_0)(p-1) + \mathrm{Var}(\gamma_1)p) + \mathrm{Cov}(\gamma_0,\gamma_1)(p(-2(\bar{x}-1)\bar{x}-1) + \bar{x}^2))}{(\bar{x}^2(\mathrm{Var}(\gamma_0) + \mathrm{Var}(\gamma_1) - 2\mathrm{Cov}(\gamma_0,\gamma_1)) + 2\bar{x}(\mathrm{Cov}(\gamma_0,\gamma_1) - \mathrm{Var}(\gamma_0)) + \mathrm{Var}(\gamma_0))^2}
$$

Thus taking the numerator and setting it equal to zero we get an FOC

$$
\bar{x} = \frac{\mathrm{Var}(\gamma_0)(p-1) + \mathrm{Var}(\gamma_1)p - 2\mathrm{Cov}(\gamma_0,\gamma_1)p \pm \sqrt{\mathrm{Var}(\gamma_0)^2(p-1)^2 + 2\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1)p(p-1) + p(\mathrm{Var}(\gamma_1)^2 p - 4\mathrm{Cov}(\gamma_0,\gamma_1)^2(p-1))}}{2(\mathrm{Var}(\gamma_0)(p-1) + \mathrm{Var}(\gamma_1)p - 2\mathrm{Cov}(\gamma_0,\gamma_1)p + \mathrm{Cov}(\gamma_0,\gamma_1))}
$$

We can see from the above solution that whenever

$$
\mathrm{Var}(\gamma_0) = 2\mathrm{Cov}(\gamma_0,\gamma_1)p = p(\mathrm{Var}(\gamma_0) + \mathrm{Var}(\gamma_1)) + \mathrm{Cov}(\gamma_0,\gamma_1)
$$

That the solution is undefined. In other words, when $\mathrm{Var}(\gamma_0) = \mathrm{Var}(\gamma_1) = \mathrm{Cov}(\gamma_0,\gamma_1) = 1$ there is no solution.

This makes sense because if the results are perfectly correlated and no group has higher variance, than any choice of representation should tell us the exact same amount about the other group. And thus there is no clear optimal solution.

Three Subgroups

First here we assume $\mathrm{Var}(\gamma_i) = 1$ and $\mathrm{Cov}(\gamma_i, \gamma_j) = 0$ for $i \neq j$
let $\beta_i = (\gamma_i - \mathrm{E}(\gamma_i)$

$$
\begin{aligned}
EJ_i = &p_1((1-c_1\bar{y}_1)\beta_1 - c_1\bar{y}_2\beta_2 - c_1\bar{y}_3\beta_3)^2 \\
&+p_2(-c_2\bar{y}_1\beta_1 + (1-c_2\bar{y}_2)\beta_2 - c_2\bar{y}_3\beta_3)^2 \\
&+p_3(-c_3\bar{y}_1\beta_1 - c_3\bar{y}_2\beta_2 + (1-c_3\bar{y}_3)\beta_3)^2
\end{aligned}
$$

$$1 - \bar{y}_1 - \bar{y}_2 = \bar{y}_3$$

Now recall that for individual i,

$$\begin{aligned} c_i &= \frac{\bar{y}_i}{\bar{y}_1^2 + \bar{y}_2^2 + \bar{y}_3^2} \quad \text{for } i \in \{1, \dots, 3\} \\ &= \frac{\bar{y}_i}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2} \quad \text{for } i \in \{1, \dots, 2\} \end{aligned}$$

also we can rewrite

$$\begin{aligned} EJ_i &= p_1(\beta_1 - c_1(\bar{y}_1\beta_1 + \bar{y}_2\beta_2 + \bar{y}_3\beta_3))^2 \\ &+ p_2(\beta_2 - c_2(\bar{y}_1\beta_1 + \bar{y}_2\beta_2 + \bar{y}_3\beta_3))^2 \\ &+ p_3(\beta_3 - c_3(\bar{y}_1\beta_1 + \bar{y}_2\beta_2 + \bar{y}_3\beta_3))^2 \\ &= p_1(\beta_1 - c_1(\bar{y}_1\beta_1 + \bar{y}_2\beta_2 + (1 - \bar{y}_1 - \bar{y}_2)\beta_3))^2 \\ &+ p_2(\beta_2 - c_2(\bar{y}_1\beta_1 + \bar{y}_2\beta_2 + (1 - \bar{y}_1 - \bar{y}_2)\beta_3))^2 \\ &+ p_3(\beta_3 - c_3(\bar{y}_1\beta_1 + \bar{y}_2\beta_2 + (1 - \bar{y}_1 - \bar{y}_2)\beta_3))^2 \end{aligned}$$

expanding c_i
We can describe these as the loss for an individual i in subgroup k

$$W_1^2 = \left(\beta_1 - \bar{y}_1 \frac{\bar{y}_1\beta_1 + \bar{y}_2\beta_2 + (1 - \bar{y}_1 - \bar{y}_2)\beta_3}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2}\right)^2$$

$$W_2^2 = \left(\beta_2 - \bar{y}_2 \frac{\bar{y}_1\beta_1 + \bar{y}_2\beta_2 + (1 - \bar{y}_1 - \bar{y}_2)\beta_3}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2}\right)^2$$

$$W_3^2 = \left(\beta_3 - (1 - \bar{y}_1 - \bar{y}_2) \frac{\bar{y}_1\beta_1 + \bar{y}_2\beta_2 + (1 - \bar{y}_1 - \bar{y}_2)\beta_3}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2}\right)^2$$

Call $\xi = \frac{\bar{y}_1\beta_1 + \bar{y}_2\beta_2 + (1 - \bar{y}_1 - \bar{y}_2)\beta_3}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2}$
Then

$$W_1^2 = (\beta_1 - \bar{y}_1\xi)^2 = \beta_1^2 - 2\beta_1\bar{y}_1\xi + \bar{y}_1^2\xi^2$$

$$W_2^2 = (\beta_2 - \bar{y}_2\xi)^2 = \beta_2^2 - 2\beta_2\bar{y}_2\xi + \bar{y}_2^2\xi^2$$

$$W_3^2 = (\beta_3 - (1 - \bar{y}_1 - \bar{y}_2)\xi)^2 = \beta_3^2 - 2\beta_3\xi(1 - \bar{y}_1 - \bar{y}_2) + \bar{y}_3^2\xi^2$$

taking expectation over γ_i we get

$$W_1^2 = \mathrm{Var}(\gamma_1) - 2\mathrm{Var}(\beta_1)\frac{\bar{y}_1^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2} + \bar{y}_1^2\xi^2$$

$$W_2^2 = \mathrm{Var}(\gamma_2) - 2\mathrm{Var}(\beta_2)\frac{\bar{y}_2^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2} + \bar{y}_2^2\xi^2$$

$$W_3^2 = \mathrm{Var}(\gamma_3) - 2\mathrm{Var}(\beta_3)\frac{(1 - \bar{y}_1 - \bar{y}_2)^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2} + (1 - \bar{y}_1 - \bar{y}_2)^2\xi^2$$

and start by assuming $\mathrm{Var}(\gamma_i) = 1$ then

$$\begin{aligned}
J =& p_1W_1^2 + p_2W_2^2 + p_3W_3^2 \\
J =& p_1[1 - 2\frac{\bar{y}_1^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2} + \bar{y}_1^2\xi^2] \\
&+ p_2[1 - 2\frac{\bar{y}_2^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2} + \bar{y}_2^2\xi^2] + \\
&(1 - p_1 - p_2)[1 - 2\frac{(1 - \bar{y}_1 - \bar{y}_2)^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2} + (1 - \bar{y}_1 - \bar{y}_2)^2\xi^2]
\end{aligned}$$

First, assume $p = p_1 = p_2 = p_3 = \frac{1}{3}$
and note that in expectation (assuming $Cov(\gamma_i, \gamma_j) = 0$ for $i \neq j$)

$$\begin{aligned}
(\bar{y}_1\beta_1 + \bar{y}_2\beta_2 + (1 - \bar{y}_1 - \bar{y}_2)\beta_3)^2 &= \mathrm{Var}(\gamma_1)\bar{y}_1^2 + \mathrm{Var}(\gamma_2)\bar{y}_2^2 + \mathrm{Var}(\gamma_3)(1 - \bar{y}_1 - \bar{y}_2)^2 \\
&= \bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2
\end{aligned}$$

In other words, in expectation

$$\xi^2 = \frac{1}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2}$$

then

$$\begin{aligned}
J =& p[3 - 2\frac{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2} + (\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2)\xi^2] \\
=& p[3 - 2 + 1] = \frac{2}{3}
\end{aligned}$$

Now for general p_1, p_2, p_3

$$\begin{aligned}
J =& p_1 W_1^2 + p_2 W_2^2 + p_3 W_3^2 \\
J =& p_1[1 - \frac{\bar{y}_1^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2}] \\
&+ p_2[1 - \frac{\bar{y}_2^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2}] + \\
&(1 - p_1 - p_2)[1 - \frac{(1 - \bar{y}_1 - \bar{y}_2)^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2}] \\
&= 1 - \frac{p_1\bar{y}_1^2 + p_2\bar{y}_2^2 + p_3(1 - \bar{y}_1 - \bar{y}_2)^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2} \\
&= \frac{(1 - p_1)\bar{y}_1^2 + (1 - p_2)\bar{y}_2^2 + (1 - p_3)(1 - \bar{y}_1 - \bar{y}_2)^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2}
\end{aligned}$$

$$\begin{aligned}
&\frac{d}{dx}\frac{(1 - p_1)x^2 + (1 - p_2)y^2 + (1 - p_1 - p_2)(1 - x - y)^2}{x^2 + y^2 + (1 - x - y)^2} = 0 \\
&\frac{2p_2x(x(y - 1) + 2y^2 - 2y + 1) - 2p_1y^2(2x + y - 1)}{(x^2 + y^2 + (1 - x - y)^2)^2} = 0 \\
&\frac{d}{dy}\frac{(1 - p_1)x^2 + (1 - p_2)y^2 + (1 - p_1 - p_2)(1 - x - y)^2}{x^2 + y^2 + (1 - x - y)^2} = 0 \\
&\frac{2p_1y(2x^2 + x(y - 2) - y + 1) - 2p_2x^2(x + 2y - 1)}{(x^2 + y^2 + (1 - x - y)^2)^2} = 0
\end{aligned}$$

We must solve the system of equations

$$2p_2x(x(y-1)+2y^2-2y+1)-2p_1y^2(2x+y-1)=0$$

$$2p_1y(2x^2+x(y-2)-y+1)-2p_2x^2(x+2y-1)=0$$

we can immediately see that $x=0$ and $y=1$ are always solutions. Likewise $(x,y)=(0,0)$ is also always a solution. We only need to investigate whether the Heissian is positive definite or not to determine which one is a minimum and which is a maximum

Let A be a vector such that

$$A=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}$$

and $\bar{A}$ be define likewise.

then it can be seen that

$$\mathrm{E}J_i = E[(A_i - c_i\bar{A})(\gamma-\bar{\gamma})]^2$$

5. Generalization allowing for k subgroups.

again we assume $\mathrm{Var}(\gamma_i)=1$ and $\mathrm{Cov}(\gamma_i,\gamma_j)=0$ for $i\neq j$

we saw from before that we can write

$$\begin{aligned} EJ_i &= p_1(\alpha_1 - c_1(\bar{y}_1\alpha_1+\bar{y}_2\alpha_2+\ldots+\bar{y}_k\alpha_k))^2 \\ &+ p_2(\alpha_2 - c_2(\bar{y}_1\alpha_1+\bar{y}_2\alpha_2+\ldots+\bar{y}_k\alpha_k))^2 \\ &\vdots \\ &+ p_k(\alpha_k - c_k(\bar{y}_1\alpha_1+\bar{y}_2\alpha_2+\ldots+\bar{y}_k\alpha_k))^2 \end{aligned}$$

expanding c_i

We can describe these as the loss for an individual i in subgroup k

$$W_1^2 = \left(\alpha_1 - \bar{y}_1 \frac{\bar{y}_1\alpha_1 + \bar{y}_2\alpha_2 + \ldots + \bar{y}_k\alpha_k}{\bar{y}_1^2 + \bar{y}_2^2 + \ldots + \bar{y}_k^2}\right)^2$$

$$W_2^2 = \left(\alpha_2 - \bar{y}_2 \frac{\bar{y}_1\alpha_1 + \bar{y}_2\alpha_2 + \ldots + \bar{y}_k\alpha_k}{\bar{y}_1^2 + \bar{y}_2^2 + \ldots + \bar{y}_k^2}\right)^2$$

$$W_3^2 = \left(\alpha_3 - \bar{y}_k \frac{\bar{y}_1\alpha_1 + \bar{y}_2\alpha_2 + \ldots + \bar{y}_k\alpha_k}{\bar{y}_1^2 + \bar{y}_2^2 + \ldots + \bar{y}_k^2}\right)^2$$

where $\bar{y}_k = (1 - \bar{x}_1 - \ldots - \bar{y}_{k-1}$
Call $\xi = \frac{\bar{y}_1\alpha_1 + \bar{y}_2\alpha_2 + \ldots + \bar{y}_k\alpha_k}{\bar{y}_1^2 + \bar{y}_2^2 + \ldots + \bar{y}_k^2}$

$$W_1^2 = (\alpha_1 - \bar{y}_1\xi)^2 = \alpha_1^2 - 2\alpha_1\bar{y}_1\xi + \bar{y}_1^2\xi^2$$

$$W_2^2 = (\alpha_2 - \bar{y}_2\xi)^2 = \alpha_2^2 - 2\alpha_2\bar{y}_2\xi + \bar{y}_2^2\xi^2$$

$$W_3^2 = (\alpha_3 - (1 - \bar{y}_1 - \bar{y}_2)\xi)^2 = \alpha_3^2 - 2\alpha_3\xi(1 - \bar{y}_1 - \bar{y}_2) + \bar{y}_3^2\xi^2$$

taking expectation over γ_i we get

$$W_1^2 = \mathrm{Var}(\gamma_1) - 2\mathrm{Var}(\gamma_1)\frac{\bar{y}_1^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2} + \bar{y}_1^2\xi^2$$

$$W_2^2 = \mathrm{Var}(\gamma_2) - 2\mathrm{Var}(\gamma_2)\frac{\bar{y}_2^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2} + \bar{y}_2^2\xi^2$$

$$W_3^2 = \mathrm{Var}(\gamma_3) - 2\mathrm{Var}(\gamma_3)\frac{(1 - \bar{y}_1 - \bar{y}_2)^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1 - \bar{y}_1 - \bar{y}_2)^2} + (1 - \bar{y}_1 - \bar{y}_2)^2\xi^2$$

and start by assuming $\mathrm{Var}(\gamma_i) = 1$ then

$$J = p_1 W_1^2 + p_2 W_2^2 + p_3 W_3^2$$

$$J = p_1[1 - 2\frac{\bar{y}_1^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1-\bar{y}_1-\bar{y}_2)^2} + \bar{y}_1^2\xi^2]$$

$$+p_2[1 - 2\frac{\bar{y}_2^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1-\bar{y}_1-\bar{y}_2)^2} + \bar{y}_2^2\xi^2]+$$

$$(1-p_1-p_2)[1 - 2\frac{(1-\bar{y}_1-\bar{y}_2)^2}{\bar{y}_1^2 + \bar{y}_2^2 + (1-\bar{y}_1-\bar{y}_2)^2} + (1-\bar{y}_1-\bar{y}_2)^2\xi^2]$$

First, assume $p = p_1 = p_2 = p_3 = \frac{1}{3}$
and note that in expectation (assuming $Cov(\gamma_i, \gamma_j) = 0$ for $i \neq j$)

$$(\bar{y}_1\alpha_1 + \bar{y}_2\alpha_2 + (1-\bar{y}_1-\bar{y}_2)\alpha_3)^2 = \mathrm{Var}(\gamma_1)\bar{y}_1^2 + \mathrm{Var}(\gamma_2)\bar{y}_2^2 + \mathrm{Var}(\gamma_3)(1-\bar{y}_1-\bar{y}_2)^2$$

$$= \bar{y}_1^2 + \bar{y}_2^2 + (1-\bar{y}_1-\bar{y}_2)^2$$

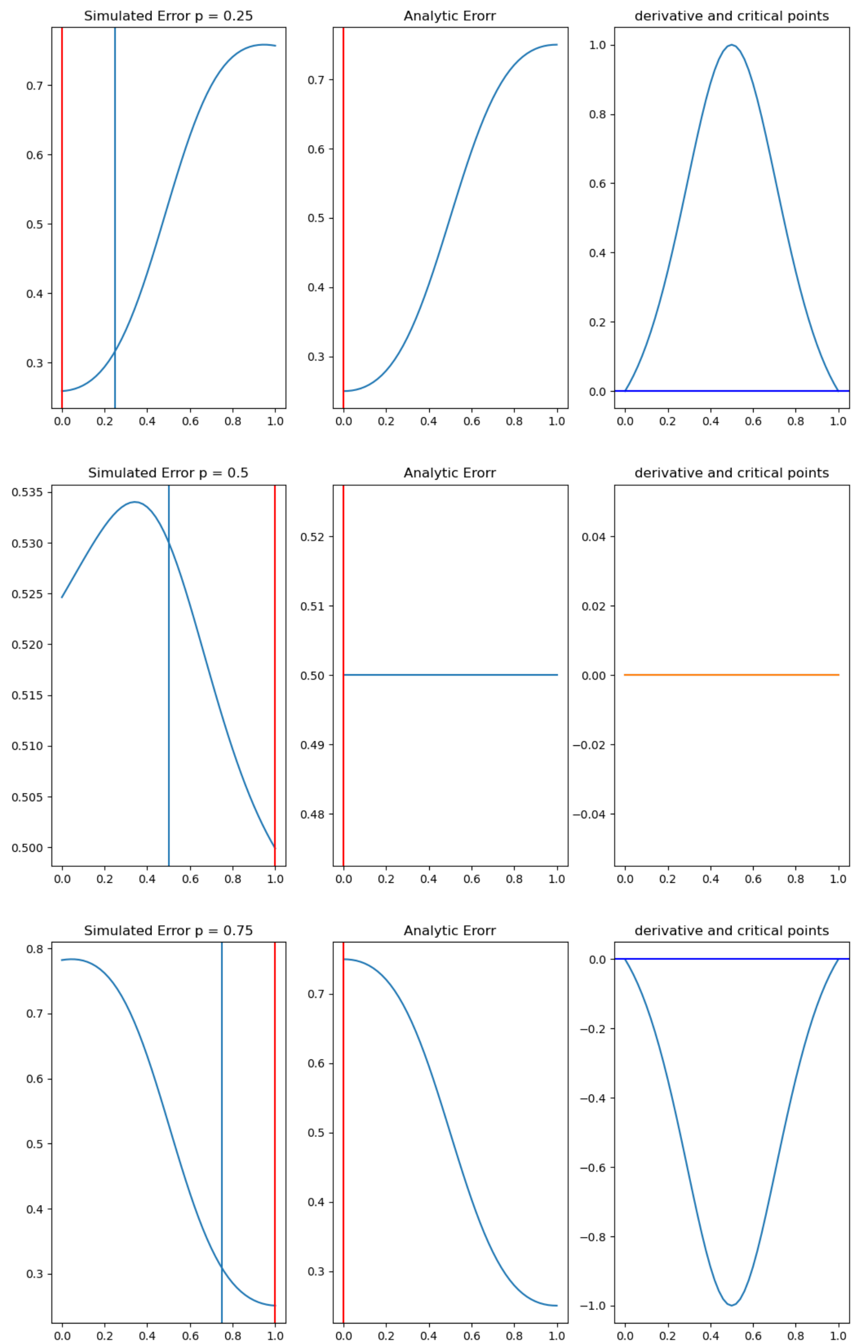

Figure 3: First column contains the simulated error for different population proportions with the population proportion shown by the blue vertical line and point of minimum error shown with the red vertical line. The second column has the analytic error computed in the math above as well as a vertical line showing the first critical point. The third column shows the first derivative with respect to $\bar{x}$.

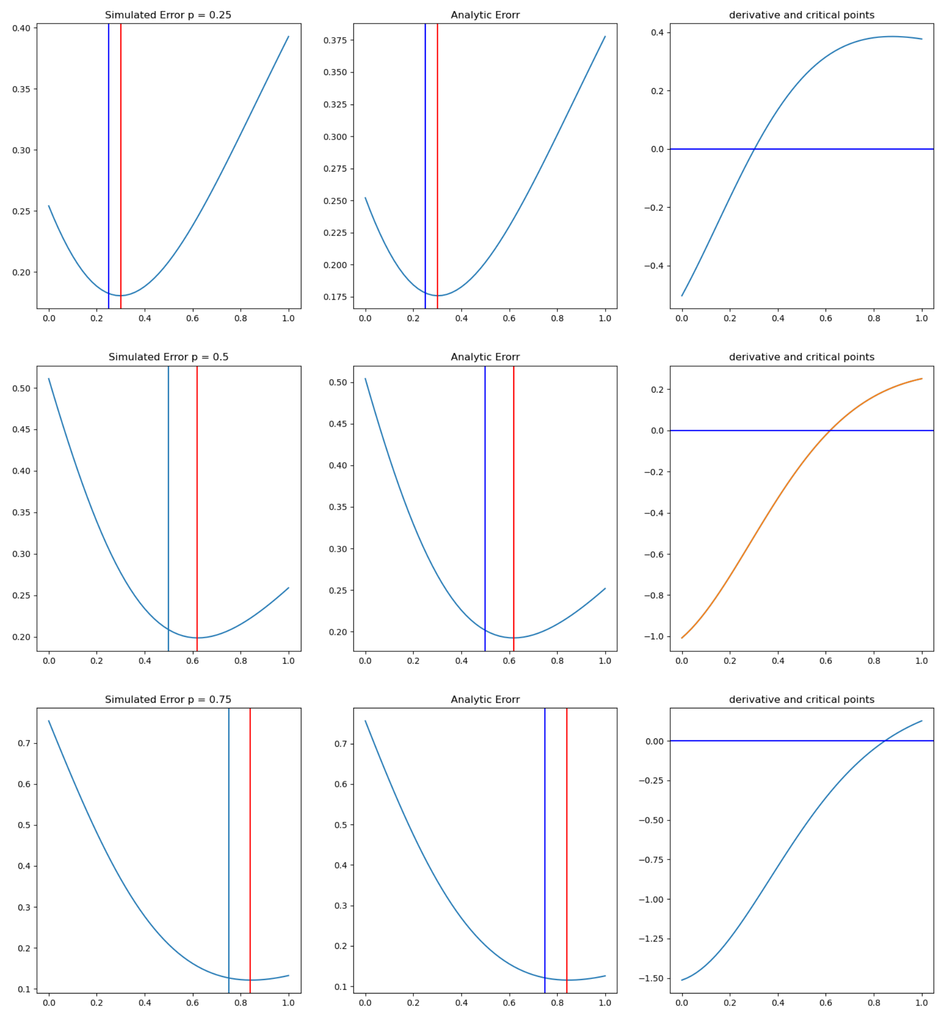

Figure 4: First column contains the simulated error for different population proportions with the population proportion shown by the blue vertical line and point of minimum error shown with the red vertical line. The second column has the analytic error computed in the math above as well as a vertical line showing the first critical point and blue line showing population proportion. The third column shows the first derivative with respect to $\bar{x}$ and where it equals zero.

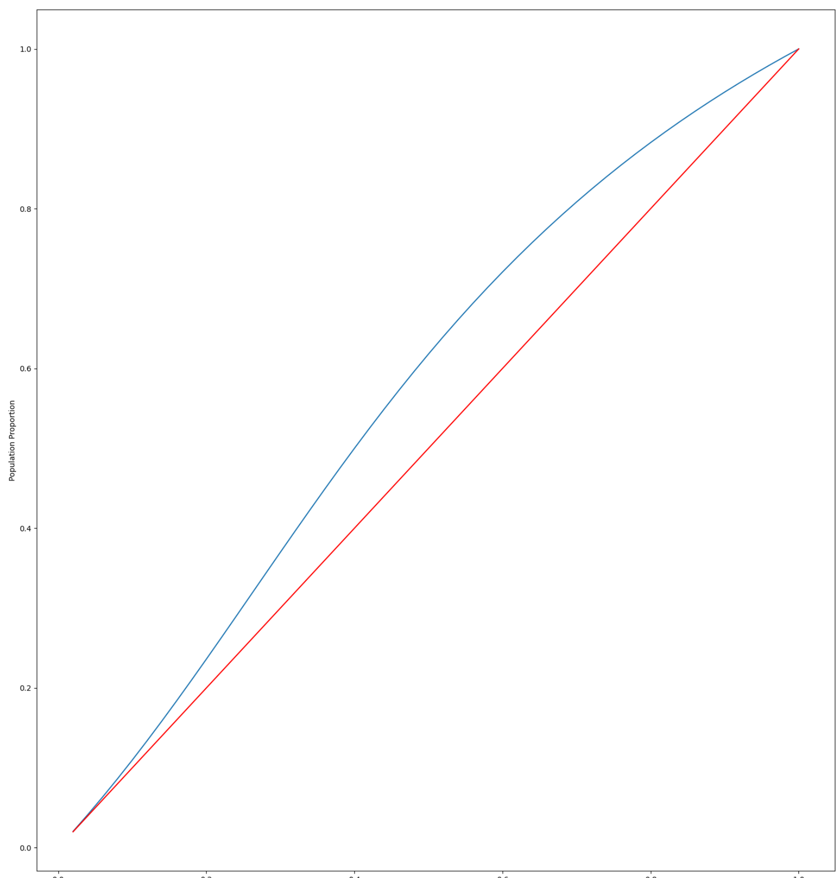

Figure 5: Here we show the optimal trial proportion of men vs the population proprtion (on the y and x axis respectively) for equal variances of 1.