Generating multilevel data
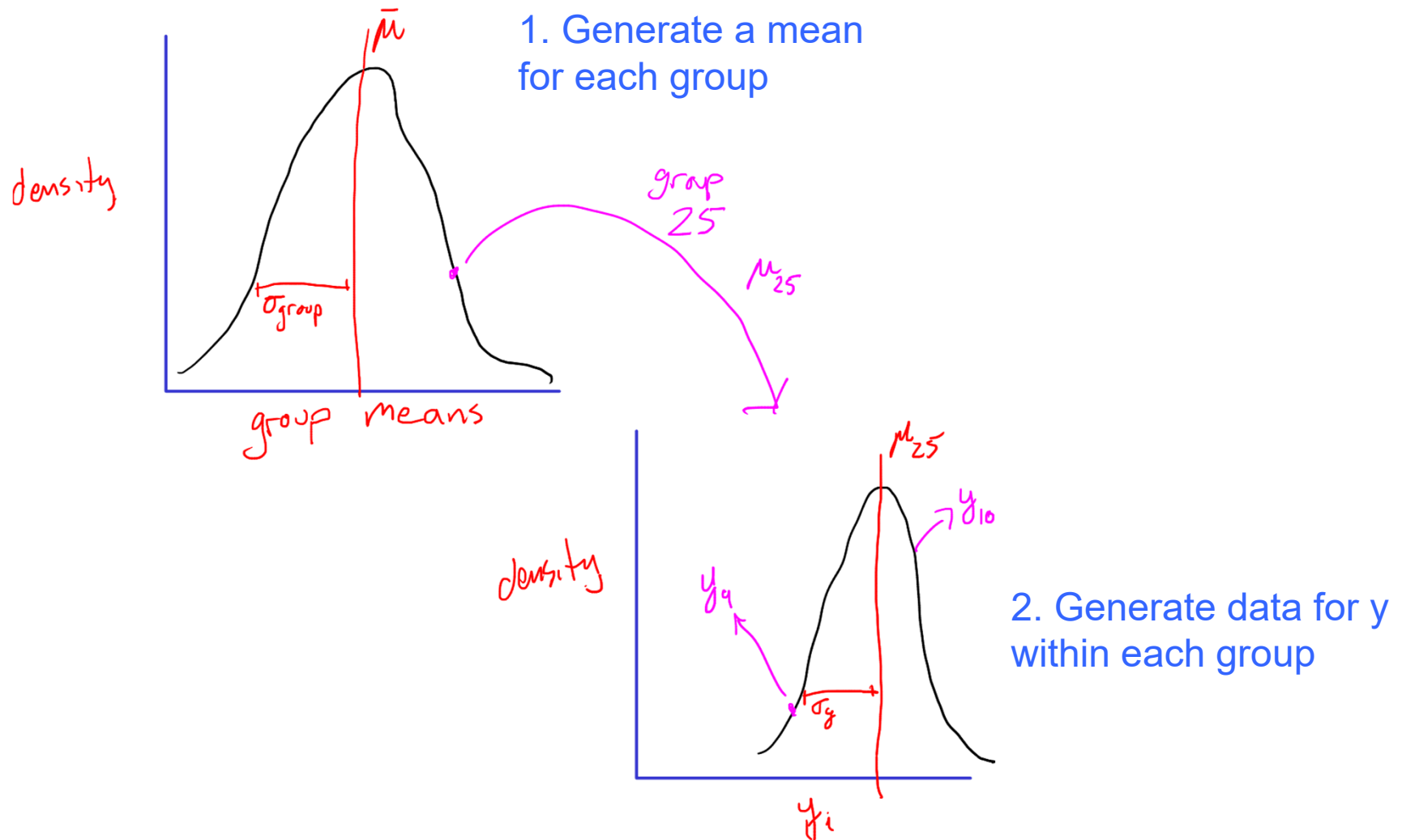

1. Generate a mean for each group

2. Generate data for y within each group

What to do when simulated data goes wrong?

Three possibilities
1. Math wrong
2. Simulation wrong
   - code error (bug) or translation of math to code
3. Fit wrong
   - training algorithm issue or model code wrong

Strategy: build from simple models
1. Make it as simple as possible at first
2. Build complexity incrementally