Appendix: Minimum and maximum for range of counterfactual household counts

## Definitions and Notation

Let:

- P = **Total population size**

- H = **Number of households**

- M_h = **Household-level mean household size**

  The **household-level** mean household size is a simple average across households as individual observations:

$$M_h = \frac{P}{H} \tag{2}$$

  For example, in Figure 1.1, the the population of $P = 4$ is divided into $H = 2$ households. Thus, the average household size is 2 individuals.

  Figure 1.1: An example population divided into two households
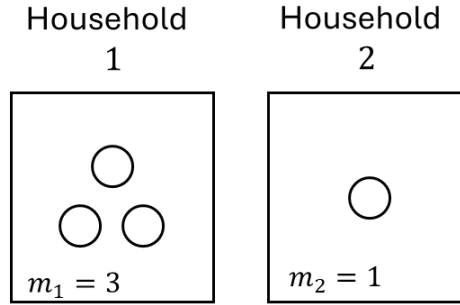
- M_p = **Person-level mean household size**

  For every person j in a population P living in a household of m_j members, the **person-level** mean household size is defined as the average size of a household that a person lives in:

$$M_p = \frac{1}{P} \sum_{j=1}^{H} m_j^2 \tag{1}$$

  For example, in Figure 1.1, three individuals live in a household of three members, and one individual lives in a household of one member. Thus, the average size of a household that a person lives in is $\frac{3+3+3+1}{4} = \frac{3^2+1^1}{4} = 2.5$ individuals.

Context

Equation (2) shows that household-level mean household size M_h and population P is adequate to deduce the number of households H. However, person-level household size M_p does not lend itself to this equivalency: the same person-level mean and population can be consistent with multiple values for the number of households.

Figure 1.2 illustrates this point in a hypothetical population with $P = 6$ and $M_p = 3$. In Scenario 1, average person-level household size is $\frac{3^2+3^2}{6} = 3$ and there are 2 households. In Scenario 2, average person-level household size is $\frac{1^2+1^2+4^2}{6} = 3$ and there are 3 households. Thus, population and person-level household size alone do not singularly determine the number of households.

Figure 1.2
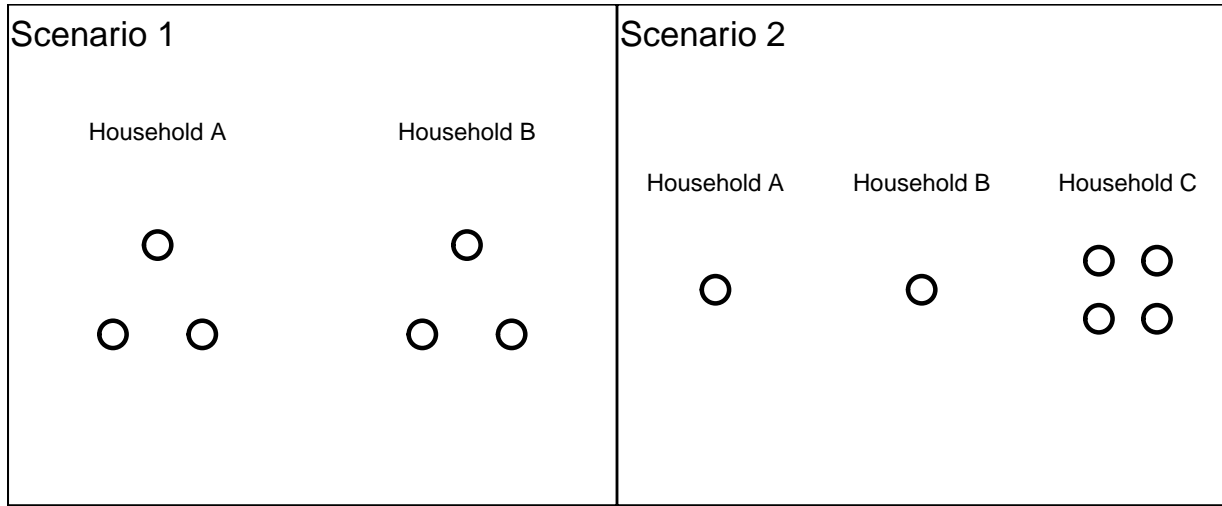
Circles represent individual people. In both scenarios, P = 6 and M$_p$ = 3, yet the number of households differs.

Assertion

There is a range of possible number of households H consistent with a known population size P and person-level average household size M_p. The range on H is bounded by a minumum of

$$H_{min} = \frac{P}{M_p} \tag{1}$$

and a maximum of

[FILL IN HERE]

Lemma 1: Minimum Number of Households

The minimum number of households consistent with a population size P and person-level average household size M_p is:

$$H_{\min} = \frac{P}{M_p} \tag{2}$$

Proof

Suppose the population P is distributed into H households with the measured household size in the jth household equal to m_j. As defined above, **person-level** mean household size is:

$$M_p = \frac{1}{P} \sum_{j=1}^{H} m_j^2 \tag{3}$$

Any configuration of P individuals across households can be transformed into any other configuration through a finite sequence of single-person transfers. The initial count of households, H, need not match the final count of households, H': we conceptualize a 'reservoir' of zero-person households to account for changes between H and H'. Forming a new household can be conceptualized as moving an individual to a zero-person household, which remains uncounted unless populated. Conversely, a household can be dissolved by moving all of its members to other households, thus turning a household with a positive number of members into a zero-person household. For example, in Figure 1.2, once can incrementally create Scenario 2 after starting from Scenario 1 through the following steps:

1. Move one individual from Household A to a newly formed Household C.
2. Move another individual from Household A to Household C.
3. Move two individuals from Household B to Household C.

Next, we calculate how the person-level mean household size changes after a single-person transfer. Suppose we begin with an arbitrary configuration of P individuals in H households, with each household m_j containing some whole number of members $m_1, m_2, \ldots, m_{H-1}, m_H$. We define a single-person transfer as a move of one individual from household $j = k$ to household $j = k + 1$. Mean household size M_p changes from:

$$M_p = \frac{1}{P} \left(\sum_{j=1}^{k-1} m_j^2 + m_k^2 + m_{k+1}^2 + \sum_{j=k+2}^{H} m_j^2 \right) \tag{4}$$

to

$$M_p' = \frac{1}{P} \left(\sum_{j=1}^{k-1} m_j'^2 + m_k'^2 + m_{k+1}'^2 + \sum_{j=k+2}^{H} m_j'^2 \right) \tag{5}$$

Person-level household size increases if $M_p' - M_p > 0$, remains the same if $M_p' - M_p = 0$, and shrinks if $M_p' - M_p < 0$. Since all household sizes $m_j = m_j'$ are unaffected by the shift, aside from the special case of $j = k$ and $j = k + 1$, the remaining terms cancel out due to equivalence.

$$M_p' - M_p = \frac{1}{P} \left(\sum_{j=1}^{k-1} m_j^2 + m_k^2 + m_{k+1}^2 + \sum_{j=k+2}^{H} m_j^2 \right) - \frac{1}{P} \left(\sum_{j=1}^{k-1} m_j'^2 + m_k'^2 + m_{k+1}'^2 + \sum_{j=k+2}^{H} m_j'^2 \right) \tag{6}$$

$$= \frac{1}{P} \left((m_k^2 + m_{k+1}^2) - (m_k'^2 + m_{k+1}'^2) \right) \tag{3}$$

$$\tag{7}$$
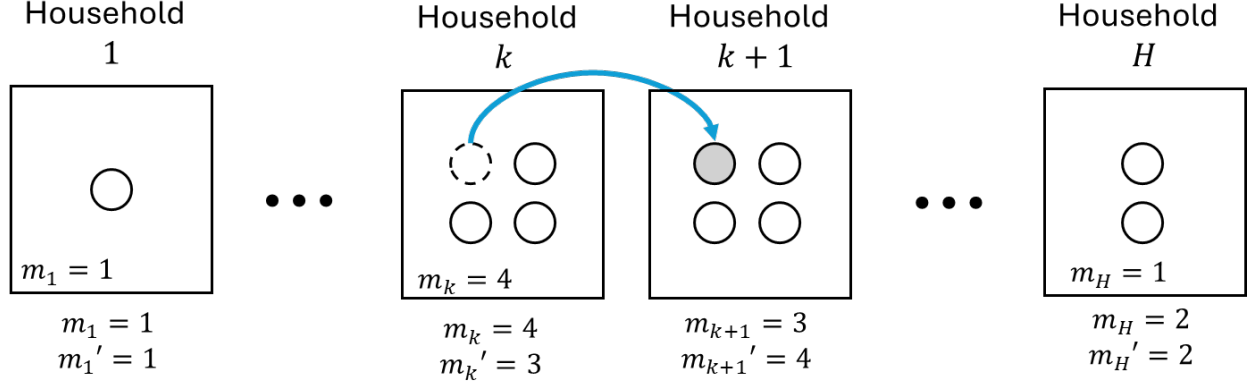
We now transition to exploring how the value $M_p' - M_p$ varies under each of the three circumstances:

1. $m_{k+1} = m_k - 1$
2. $m_{k+1} < m_k - 1$
3. $m_{k+1} > m_k - 1$

Since a single-person transfer involves a move of one person from household k to household $k+1$, the post-transfer number of household members in household k is $m'_k = m_k - 1$ while the post-transfer number of household members in household $k+1$ is $m'_{k+1} = m_{k+1} + 1$.

Case 1: a symmetric transfer

Figure 1.3: An example single-person transfer where $m_{k+1} = m_k - 1$

Household 1		Household k	Household $k+1$		Household H

$m_1 = 1$ · · · $m_k = 4$ $m_{k+1} = 3$ · · · $m_H = 1$

$m_1 = 1$	$m_k = 4$	$m_{k+1} = 3$		$m_H = 2$
$m_1' = 1$	$m_k' = 3$	$m_{k+1}' = 4$		$m_H' = 2$

Initially, household h has $m_k = 4$ members and household $k+1$, has $m_{k+1} = 3$ members. One individual is moved from household k to household $k+1$, resulting in $m'_k = 3$ members in household k and $m'_{k+1} = 4$ members in household $k+1$.

In Case 1, the overall person-level mean household size remains unchanged after the shift, because the two affected households have precisely a one-person difference in their sizes. When one person moves from the larger household to the smaller one, the decrease in the squared size of the original larger household is exactly offset by the increase in the squared size of the receiving household.

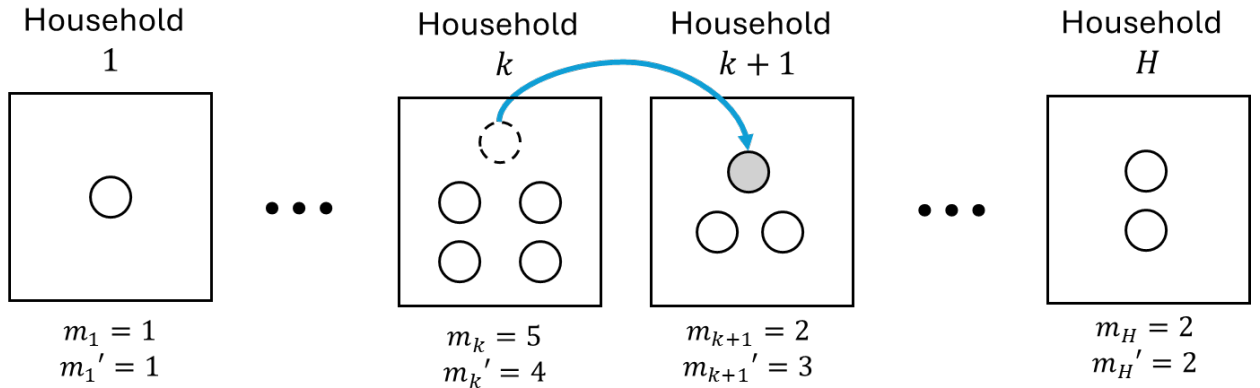$$M'_p - M_p = \frac{1}{P}\left((m_k^2 + m_{k+1}^2) - (m_k'^2 + m_{k+1}'^2) \right) \tag{8}$$

$$= \frac{1}{P}\left((m_k^2 + (m_k - 1)^2) - ((m_k - 1)^2 + m_k^2) \right) \tag{9}$$

$$= 0 \tag{10}$$

$$\tag{11}$$

Case 2: a transfer from a larger household to a smaller household

Figure 1.4: An example single-person transfer where $m_{k+1} < m_k - 1$

Household 1		Household k	Household $k+1$		Household H

· · · · · ·

$m_1 = 1$	$m_k = 5$	$m_{k+1} = 2$		$m_H = 2$
$m_1' = 1$	$m_k' = 4$	$m_{k+1}' = 3$		$m_H' = 2$

Initially, household h has $m_k = 5$ members and household $k+1$, has $m_{k+1} = 2$ members. One individual is moved from household k to household $k+1$, resulting in $m'_k = 4$ members in household k and $m'_{k+1} = 3$ members in household $k+1$.
[ADD IN EXPLANATION ON HOW THIS TRANSFER REDUCES AVERAGE HOUSEHOLD SIZE]

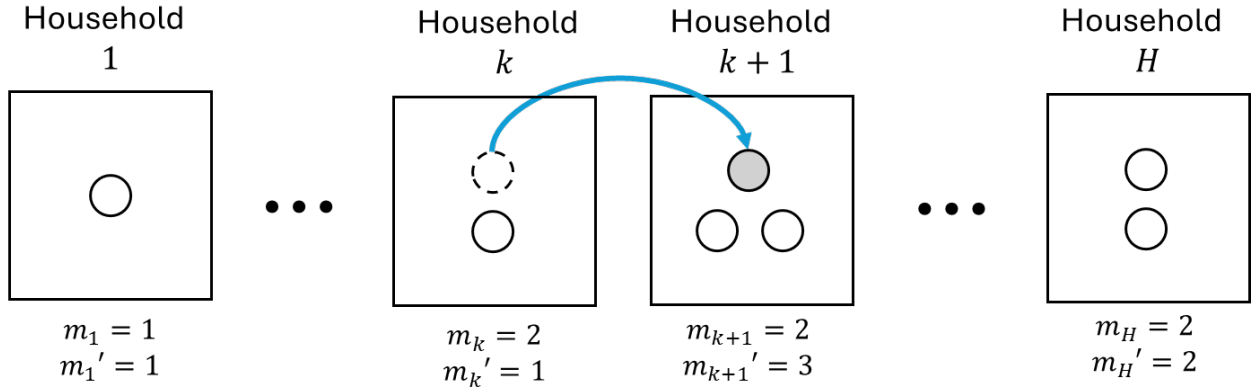$$M'_p - M_p = \frac{1}{P}\left((m_k^2 + m_{k+1}^2) - (m'^2_k + m'^2_{k+1})\right) \tag{12}$$

$$= \frac{1}{P}\left(2m_k^2 - 2m_k + 1\right) - (m) \qquad = \frac{1}{P}\left((m_k^2 + (m_k-1)^2) - ((m_k-1)^2 + m_k^2)\right) \tag{13}$$

$$= 0 \tag{14}$$

$$\tag{15}$$

Case 3: a transfer from a smaller household to a larger (or equal-size) household

Figure 1.4: An example single-person transfer where $m_{k+1} > m_k - 1$

Initially, household h has $m_k = 2$ members and household $k+1$, has $m_{k+1} = 2$ members. One individual is moved from household k to household $k+1$, resulting in $m'_k = 1$ member in household k and $m'_{k+1} = 3$ members in household $k+1$.
[ADD IN EXPLANATION ON HOW THIS TRANSFER INCREASES AVERAGE HOUSEHOLD SIZE]

[NEXT STEPS: Outline below, to-do more thoroughly]
After cases have been discussed, demonstrate that a partition of a population into equally-sized households, each with $m_j = M_p$, will result in the minimal H consistent with M_p. Any attempt to further reduce number of households would require "dissolving" an existing household and transferring all m_j members to other household(s). But this would necessitate "Case 3" transfers, which necessarily increase M_p, as shown above. Thus, it is impossible to formulate a configuration consistent with P and M_p that has any fewer than H households.
[NOTE: Above explanation is only consistent when M_p divides cleanly into P an integer number of times, i. The proof can be extended to cases where i is not an integer by invoking Case 1 transfers and arguing that the lowest housing configuration consistent with M_p and P is a set of households where all $m_j \in \{\lfloor M_p \rfloor, \lceil M_p \rceil\}$. And that this exception doesn't matter anyway because the number of H that can be constructed using $m_j \in \{\lfloor M_p \rfloor, \lceil M_p \rceil\}$ such that the population sums to P is always going to be greater than the value P/M_p.]

Lemma 2: Maximum Number of Households

Lemma:
The maximum number of households consistent with a population size P and person-level average household

size M_p is:

$$TBD \tag{16}$$

Proof:
This is more of an accessory proof, since the maximum bounds on H are so large as to not be useful in practical contexts. However, in this section, we will demonstrate that the sparsest configuration of households is given in a case where you produce some number of 1-person households and then put the rest of the population into one remaining large household. This proof will rely on Case 2 transfers, described above.

---

conclusion

Reiterate the minima and maxima, take it back to the example at the begining and show that the two configurations listed in Figure 1.2 are actually the two special cases of minima and maxima.