

# Motivation: optimal sample size allocation

Week 3 (3.4)

Stat 260, St. Clair

# Tradeoff: Cost vs. Precision

As  $n$  (sample size) increases:

- SE's get decrease (more precise) but
- sampling costs increase

# SRS example

- $N = 3000$  units
- Assume  $S = 1$  for our measurement of interest

**Cost:** costs per unit is  $c = \$2$

$$\text{total cost} = C(n) = \$2n$$

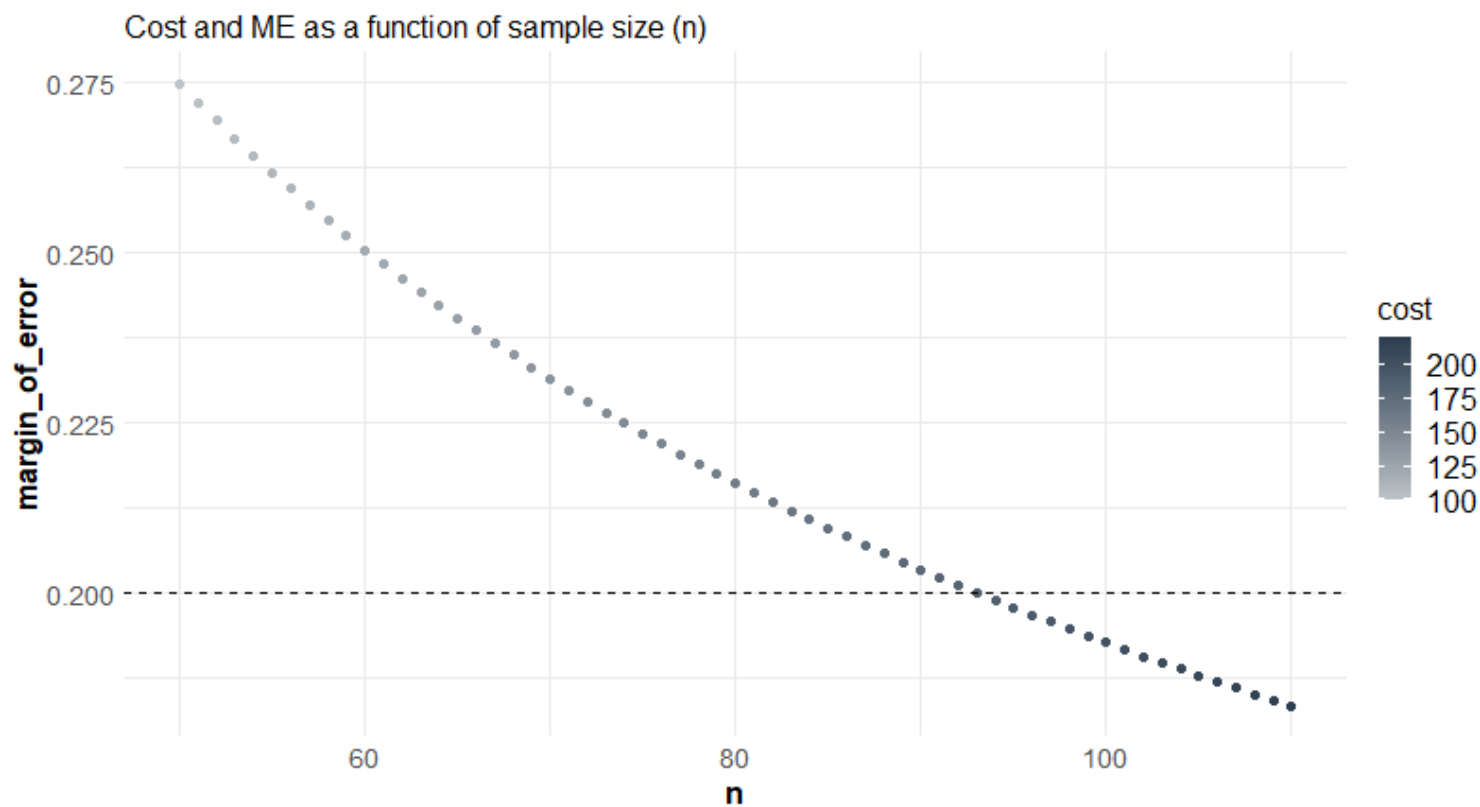
**Precision:** 95% margin of error for estimating the mean

$$ME(n) = 1.96 \times SE(\bar{y}_{srs}) = 1.96 \times \sqrt{\left(1 - \frac{n}{3000}\right) \frac{1}{n}}$$

# SRS example: determine the $n$ that...

**Constraint:** ME of at most 0.2

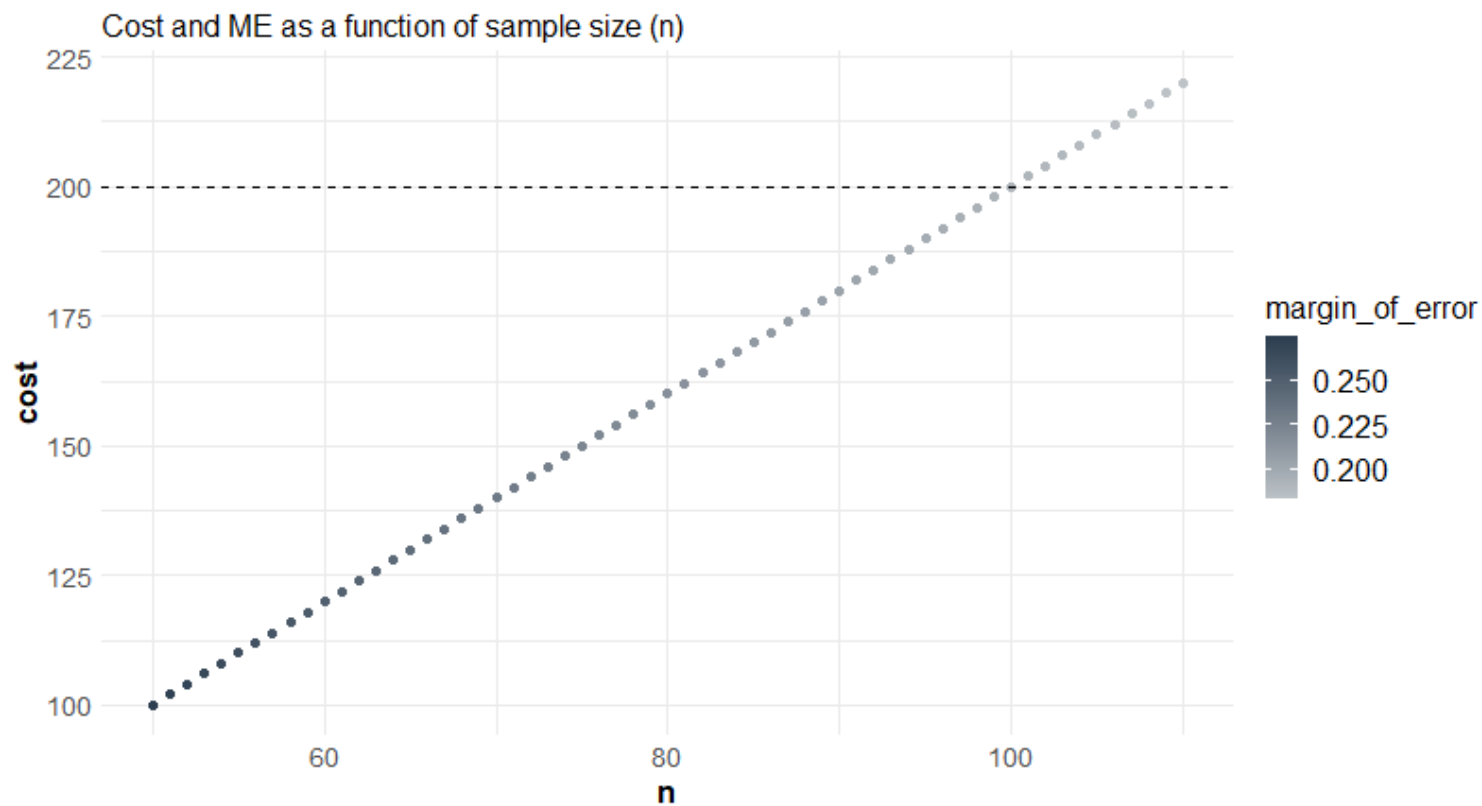
**Optimize:** minimize cost under this constraint



# SRS example: determine the $n$ that...

**Constraint:** Costs of at most \$200

**Optimize:** minimize margin of error (SE) under this constraint



# Stratified problem:

Issue: **Both** costs and precision can depend on how we **allocate** our overall sample size to each stratum

- Strata may be more/less costly to sample
- Measurements within stratum may have different SDs  $S_h$
- The **allocation** fraction for stratum  $h$  is

$$a_h = \frac{n_h}{n}$$

- Must have  $\sum_{h=1}^H a_h = 1$

# Stratified example

- $H = 3$  strata with  $N_h = 1000$  and  $S_h = 1$

**Cost:** costs per unit in each stratum are  $c_1 = 1, c_2 = 2, c_3 = 3$

$$\text{total cost} = C(n, a_1, a_2) = \$1a_1n + \$2a_2n + \$3(1 - a_1 - a_2)n$$

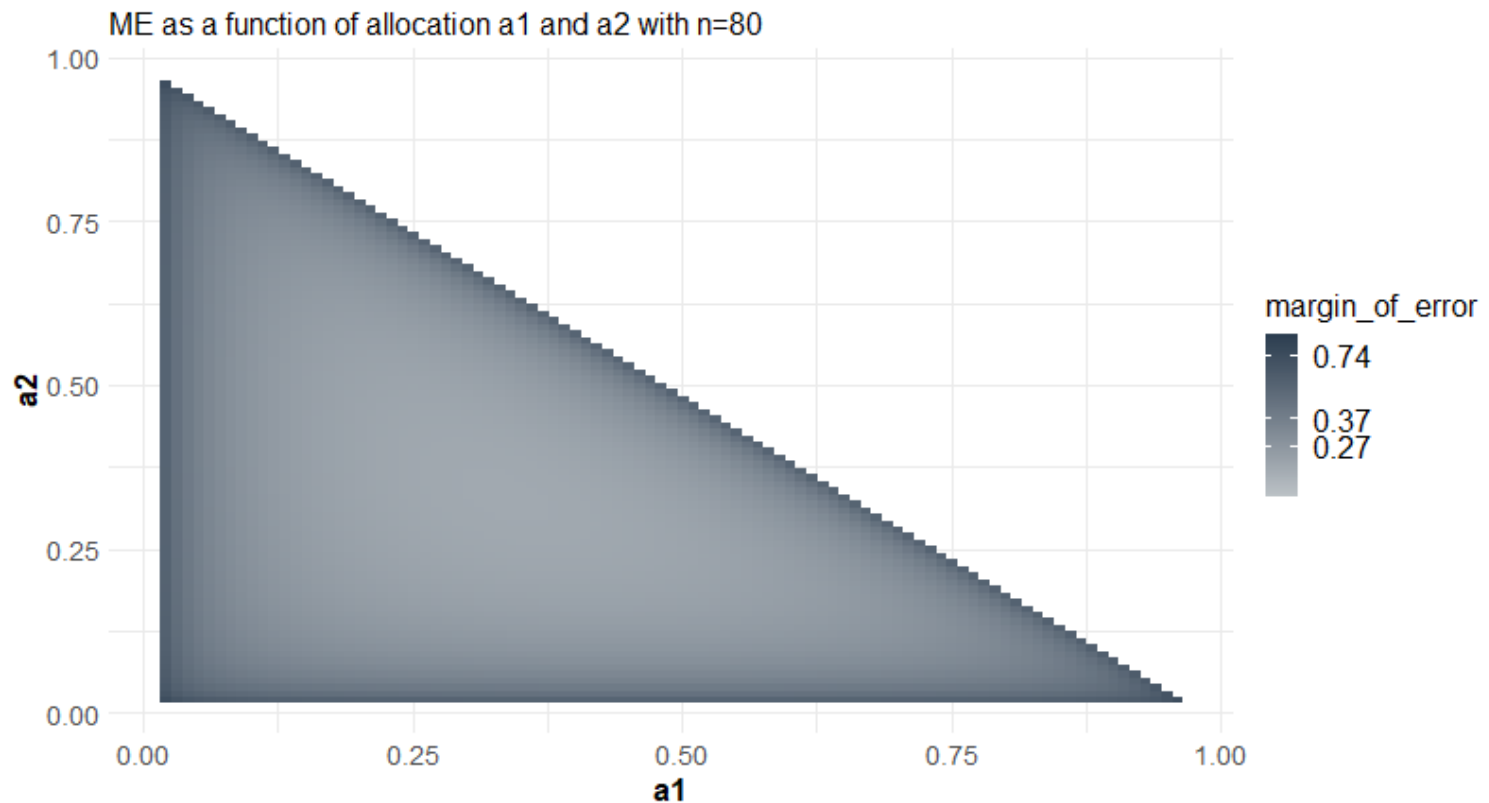
**Precision:** 95% margin of error for estimating the mean

$$ME(n, a_1, a_2) = 1.96 \times \sqrt{\sum_{h=1}^3 \left( \frac{1000}{3000} \right)^2 \left( 1 - \frac{a_h n}{1000} \right) \frac{1}{a_h n}}$$

# Stratified example: determine the $n$ , $a_1$ , $a_2$ that...

**Constraint:** Costs equal to \$200

**Optimize:** minimize margin of error (SE) under this constraint





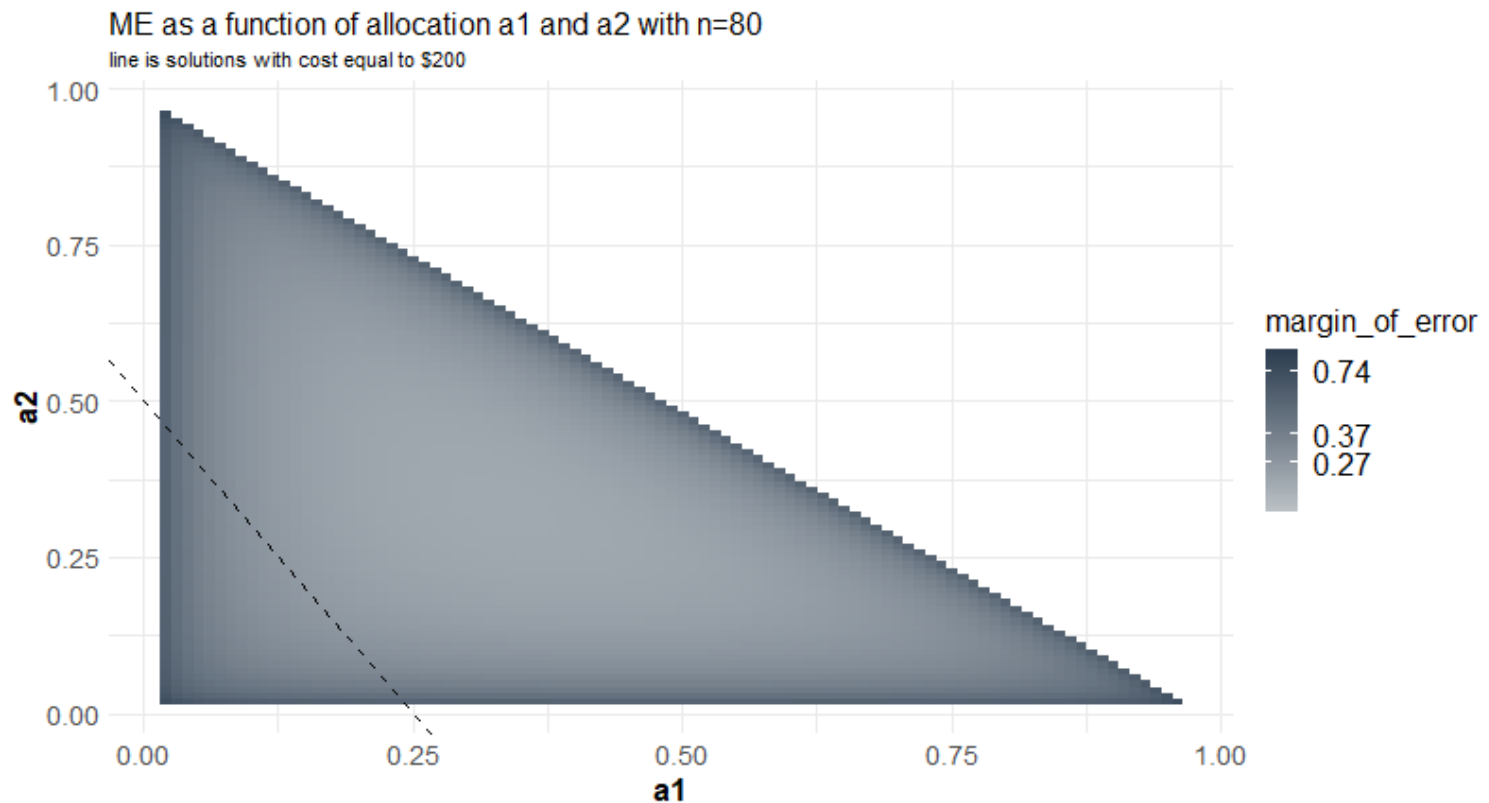
Stratified example: determine the  $n$ ,  $a_1$ ,  $a_2$  that...

**Constraint:** Costs equal to \$200

# Stratified example: determine the $n$ , $a_1$ , $a_2$ that...

**Constraint:** Costs equal to \$200

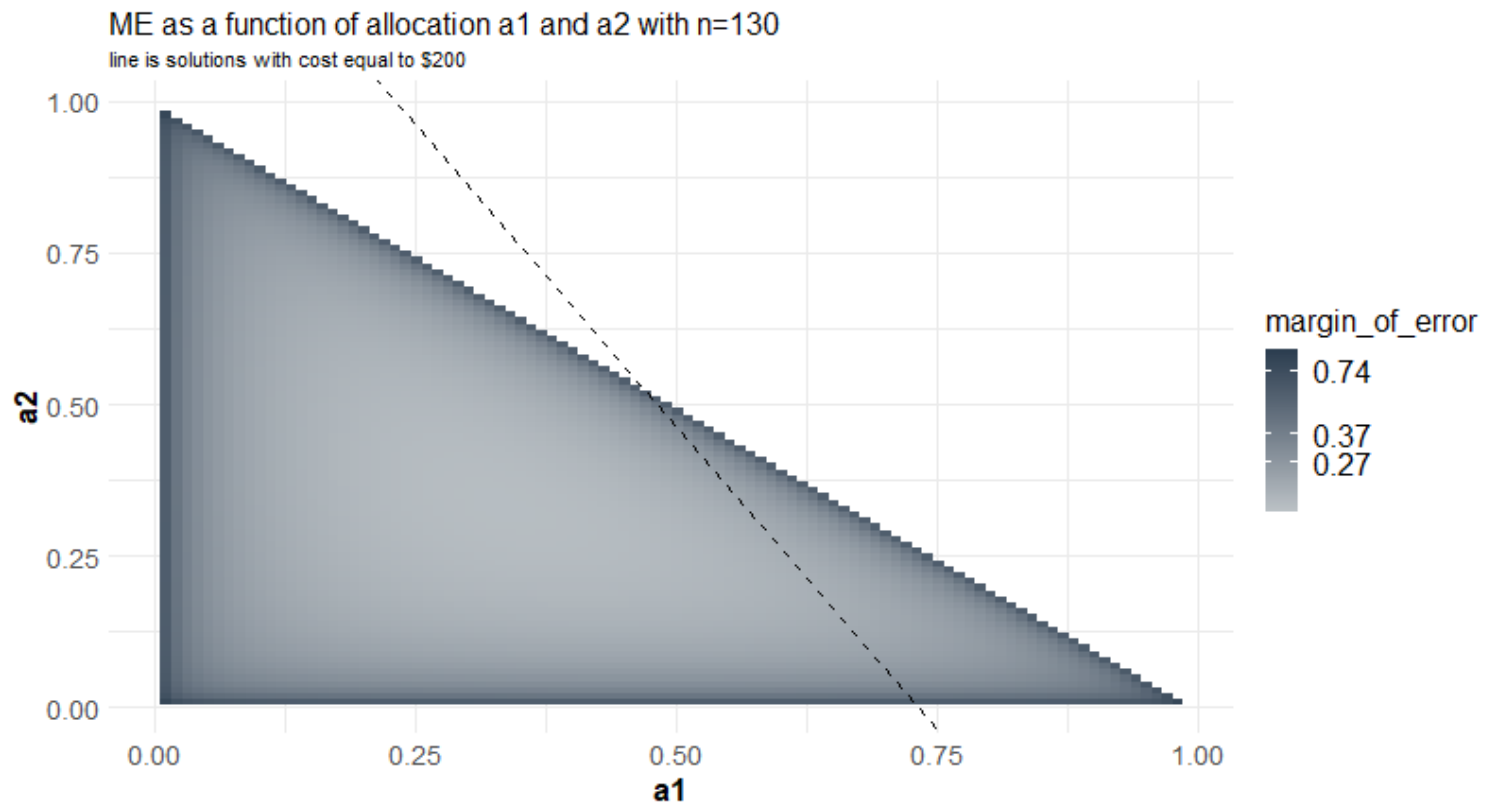
**Optimize:** minimize margin of error (SE) under this constraint



# Stratified example: determine the $n$ , $a_1$ , $a_2$ that...

**Constraint:** Costs equal to \$200

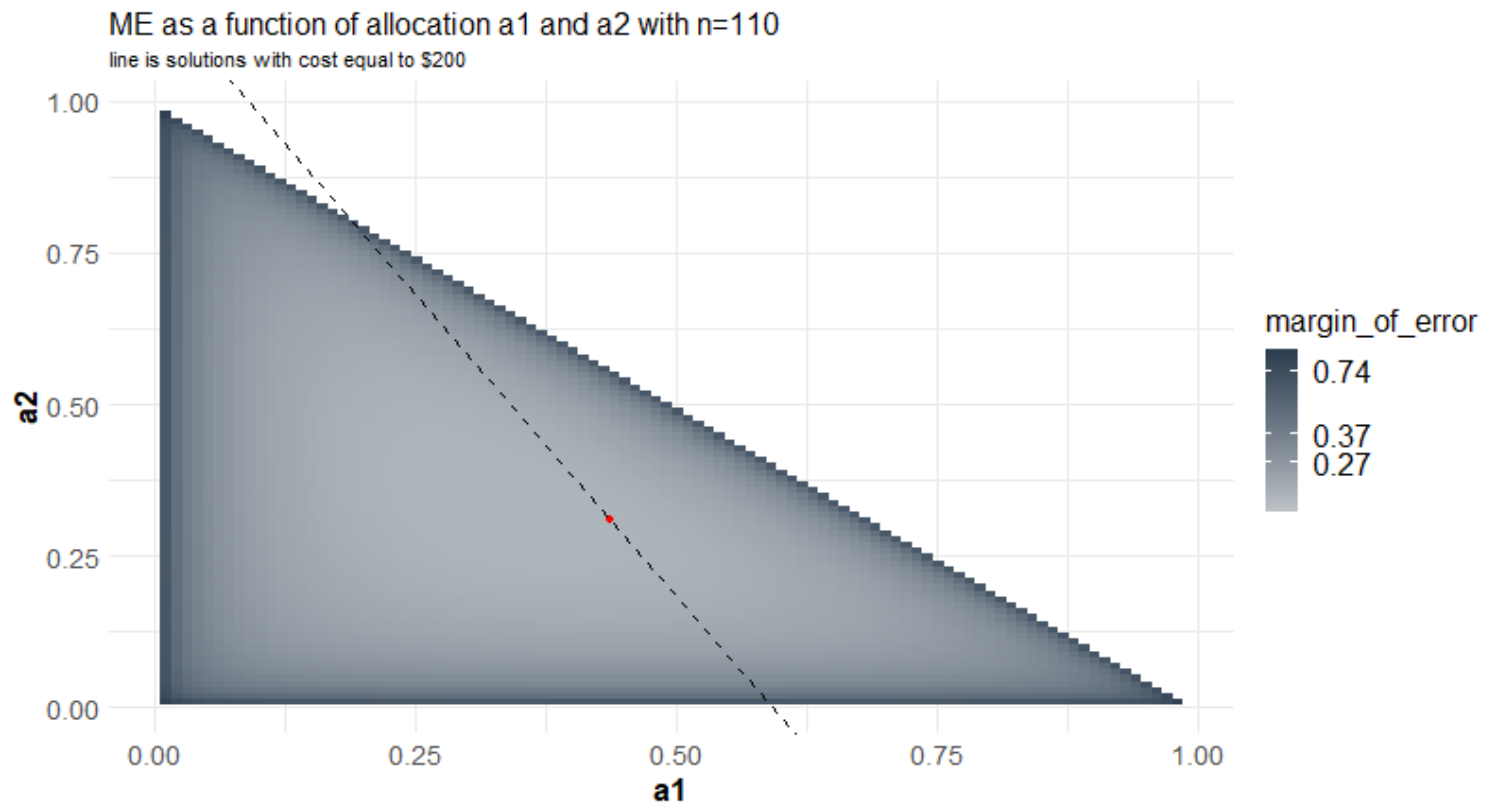
**Optimize:** minimize margin of error (SE) under this constraint



# Stratified example: determine the $n$ , $a_1$ , $a_2$ that...

**Constraint:** Costs equal to \$200

**Optimize:** minimize margin of error (SE) under this constraint



Stratified example: determine the  $n$ ,  $a_1$ ,  $a_2$  that...

**Constraint:** Costs equal to \$200

**Optimize:** minimize margin of error (SE) under this constraint

