**Part B. Q2**

**Methods**

A sample size calculation is conducted to detect an increase of at least 250 grams in mean infant birthweight associated with the smoking cessation intervention program for women subjects from the *1998* *National Maternal and Infant Health Survey* (*NMIHS*) dataset who are heavy smokers (defined as ≥10 cigarettes per day before and during the first trimester of pregnancy). The calculation formula used is shown below, where is the minimum relevant difference.

The data used for this analysis is imported as a subset of the NMIHS dataset, including only women who are heavy smokers. Missing values are then deleted for better accuracy of results.

**Assumptions and Parameters**

Certain assumptions and parameters need to be specified for this analysis. First, the outcome of interest, infant birthweight, is continuous and normally distributed within both the smoking intervention group and the control group.

The null and alternative hypotheses are H0: μintervention = μcontrol and H1: μintervention μcontrol. A 2-samplet-test will be used to test the hypothesis. There will be random and equal allocation of subjects to either the smoking intervention group or the control group, so that approximately N/2 subjects will be assigned to the intervention group and N/2 to the control group.

The expected mean infant birthweight in the control group is μcontrol = 2919.70 grams. An expected effect of interest is for smoking intervention to increase infant birthweight by 250 grams, so the expected mean birthweight in the intervention group is μintervention = 3169.7 grams. In other words, the minimum relevant difference Δ = 250 grams.

A significance level of = 0.05 is used for the 1-sided alternative hypothesis test. A power of 90% for detecting an increase of at least 250 grams in mean infant birthweight from the control group compared to the intervention group.

**Results**

Using the standard deviation of birthweight at = 698.50, the required sample size value is 267.41, rounding up to N = 268. Random assignment will allocate 134 women to the intervention group and 134 women to the control group.

**Code**

**data** b2;

set library.nmihs88\_train\_278;

if smoking = **3**;

**run**;

**proc** **means** data = b2 NMISS N;

**run**;

**data** b2;

set b2;

if parity ne **.** AND bwt ne **.** AND male ne **.**;

**run**;

**proc** **means** data = b2 NMISS N mean std;

**run**;

**data** calc\_n;

delta = **250**;

sigma = **698.50**;

alpha = **0.05**;

sides = **1**;

power = **0.90**;

Qe = **0.5**;

zalpha = probit(**1**-alpha/sides);

zbeta = probit(power);

Qc = **1** - Qe;

N = (sigma\*\***2**\*(**1**/Qe + **1**/Qc)\*(zalpha + zbeta)\*\***2**)/delta\*\***2**;

**run**;

**proc** **print** data = calc\_n;

**run**;

**Output**

**Table

Description automatically generated**

**Table, calendar

Description automatically generated**