Stat 628: Data Science Practicum

Module 2: Goals and Data Background

Goal:

Accurate measurement of body fat is inconvenient/costly and it is desirable to have easy methods of estimating body fat that are not inconvenient/costly. In this project, your group will come up with a **simple, robust, and accurate model (i.e. “rule-of-thumb”)** to estimate percentage of body fat using clinically available measurements. Your “rule-of-thumb” will be based on a real data set of 252 men with measurements of their percentage of body fat and various body circumference measurements.

You’ll have to make some trade-offs between simplicity, robustness, and accuracy in your final model/rule-of-thumb and there is no single, correct answer. Instead, we are interested in:

1. How you made this trade-off
2. How you justify this trade-off with statistical analysis or relevant background information
3. Whether your statistical and non-statistical arguments are correct and convincing

Data Background:

A variety of popular health books suggest that the readers assess their health, at least in part, by estimating their percentage of body fat. In Bailey (1994), for instance, the reader can estimate body fat from tables using their age and various skin-fold measurements obtained by using a caliper. Other texts give predictive equations for body fat using body circumference measurements (e.g. abdominal circumference) and/or skin-fold measurements. See, for instance, Behnke and Wilmore (1974), pp. 66-67; Wilmore (1976), p. 247; or Katch and McArdle (1977), pp. 120-132).

Percentage of body fat for an individual can be estimated once body density has been determined. Folks (e.g. Siri (1956)) assume that the body consists of two components - lean body tissue and fat tissue. Letting

D = Body Density (gm/cm^3)

A = proportion of lean body tissue

B = proportion of fat tissue (A+B=1)

a = density of lean body tissue (gm/cm^3)

b = density of fat tissue (gm/cm^3)

we have

D = 1/[(A/a) + (B/b)]

solving for B we find

B = (1/D)\*[ab/(a-b)] - [b/(a-b)].

Using the estimates a=1.10 gm/cm^3 and b=0.90 gm/cm^3 (see Katch and McArdle (1977), p. 111 or Wilmore (1976), p. 123) we come up with "Siri's equation":

Percentage of Body Fat (i.e. 100\*B) = 495/D - 450.

Volume, and hence body density, can be accurately measured a variety of ways. The technique of underwater weighing "computes body volume as the difference between body weight measured in air and weight measured during water submersion. In other words, body volume is equal to the loss of weight in water with the appropriate temperature correction for the water's density" (Katch and McArdle (1977), p. 113). Using this technique,

Body Density = WA/[(WA-WW)/c.f. - LV]

where (Katch and McArdle (1977), p. 115)

WA = Weight in air (kg)

WW = Weight in water (kg)

c.f. = Water correction factor (=1 at 39.2 deg F as one-gram of water occupies exactly one cm^3 at this temperature, =.997 at 76-78 deg F)

LV = Residual Lung Volume (liters)

Other methods of determining body volume are given in Behnke and Wilmore (1974), p. 22 ff.

Unfortunately, the above process of determining body volume by underwater submersion, while accurate, can be cumbersome and difficult to use by doctors who want to and easily quickly determine a patient’s body fat percentage based on commonly available measurements, even if it means sacrificing some accuracy guaranteed by underwater measurements.

The commonly available measurements include age, weight, height, bmi, and various body circumference measurements. In particular, the variables listed below (from left to right in the data set) are:

ID number of individual

Percent body fat from Siri's (1956) equation

Density determined from underwater weighing

Age (years)

Weight (lbs)

Height (inches)

Adioposity (bmi)

Neck circumference (cm)

Chest circumference (cm)

Abdomen circumference (cm)

Hip circumference (cm)

Thigh circumference (cm)

Knee circumference (cm)

Ankle circumference (cm)

Biceps (extended) circumference (cm)

Forearm circumference (cm)

Wrist circumference (cm)

Measurement standards are listed in Benhke and Wilmore (1974), pp. 45-48 where, for instance, the abdomen 2 circumference is measured "laterally, at the level of the iliac crests, and anteriorly, at the umbilicus."

References:

Bailey, Covert (1994). \_Smart Exercise: Burning Fat, Getting Fit\_, Houghton-Mifflin Co., Boston, pp. 179-186.

Behnke, A.R. and Wilmore, J.H. (1974). \_Evaluation and Regulation of Body Build and Composition\_, Prentice-Hall, Englewood Cliffs, N.J.

Siri, W.E. (1956), "Gross composition of the body", in \_Advances in Biological and Medical Physics\_, vol. IV, edited by J.H. Lawrence and C.A. Tobias, Academic Press, Inc., New York.

Katch, Frank and McArdle, William (1977). \_Nutrition, Weight Control, and Exercise\_, Houghton Mifflin Co., Boston.

Wilmore, Jack (1976). \_Athletic Training and Physical Fitness: Physiological Principles of the Conditioning Process\_, Allyn and Bacon, Inc., Boston.