

Pulse Dataset Information

Link: <http://www.statsci.org/data/oz/ms212.html>

Keywords: Paired comparisons, ANOVA, multiple regression, binomial proportions

Source: The data was supplied by Dr Richard J. Wilson, Department of Mathematics, University of Queensland. Original data file is tab-delimited text.

Description:

110 students in an introductory statistics class (MS212 taught by Professor John Eccleston and Dr Richard Wilson at The University of Queensland) participated in a simple experiment. The students took their own pulse rate. They were then asked to flip a coin. If the coin came up heads, they were to run in place for one minute. Otherwise they sat for one minute. Then everyone took their pulse again. The pulse rates and other physiological and lifestyle data are given in the data. There was missing data for one student and seemingly incorrect values for heights for two students. These observations were removed resulting in **107 subjects in the final dataset**.

Five class groups between 1993 and 1998 participated in the experiment. The lecturer, Richard Wilson, was concerned that some students would choose the less strenuous option of sitting rather than running even if their coin came up heads, so in the years 1995-1998 a different method of random assignment was used. In these years, data forms were handed out to the class before the experiment. The forms were pre-assigned to either running or non-running and there were an equal number of each. In 1995 and 1998 not all of the forms were returned so the numbers running and sitting was still not entirely controlled.

Variable	Description	Summary Statistics				
HEIGHT	Height (cm)	Mean = 173.3 cm		Std Dev = 10.17		
WEIGHT	Weight (kg)	Mean = 66.8 kg		Std Dev = 14.88		
AGE	Age (years)	Mean = 20.6 yrs		Std Dev = 3.97		
GENDER	Sex (1 = Male, 2 =Female)	Males: 54.2%		Females: 45.8%		
SMOKES	Regular smoker? (1 = Yes, 2 = No)	Yes: 10.3%		No: 89.7%		
ALCOHOL	Regular drinker? (1 = Yes, 2 = No)	Yes: 62.6%		No: 37.4%		
EXERCISE	Frequency of exercise (1 = High, 2 = Moderate, 3 = Low)	High: 13.1%	Mod: 53.3%	Low: 33.6%		
TRT	Whether the student ran or sat (1 = Ran, 2 = Sat)	Ran: 42.1%		Sat: 57.9%		
PULSE1	First pulse measurement (rate per minute)	Mean = 75.2		Std Dev = 12.68		
PULSE2	Second pulse measurement (rate per minute)	Mean = 96.2		Std Dev = 31.55		
YEAR	Year of class (93 - 98) – values are counts ⇒	93: 26	95: 22	96: 21	97: 22	98: 16

Analysis:

- The paired comparison between Pulse1 and Pulse2 should depend heavily on Ran.
- Is there evidence that some students didn't run even though their coin toss came up heads?
- How does Pulse1 depend on the lifestyle and physiological measurements? Are frequent exercisers fitter?
- How does the Pulse2-Pulse1 difference depend on lifestyle and physiological measurements? Do frequent exercisers run more energetically?
- Are there discernable differences between the years?

New Variable Information

We will create a few new variables for this dataset throughout the tutorials. Some information about these variables is given below.

Body Mass Index Calculation(kg/(m²))

For our data, the weight is measured in kg but the height is measured in cm. We must divide the height by 100 to get the height in meters. We will also round the values to 1 decimal place. The 0.1 in the round function “Rnd” tells SPSS to provide the value rounded to the nearest 0.1 or 1 decimal place.

$$\text{Rnd}(\text{BMI} = \text{WEIGHT}/(\text{HEIGHT}/100*\text{HEIGHT}/100), 0.1)$$

BMI Groups: We will create a categorized version of BMI along with binary versions for Obese vs not Obese and Overweight (or obese) vs. not Overweight (or obese).

1 = Underweight	(0 ≤ BMI < 18.5)
2 = Normal	(18.5 ≤ BMI < 25)
3 = Overweight	(25 ≤ BMI < 30)
4 = Obese	(BMI ≥ 30)

Natural Logarithm Transformation

We will create the log transformation for WEIGHT and possibly for BMI.

$$\text{In SPSS: NEWVAR} = \text{LN}(\text{OLDVAR})$$

Standardize a Variable

We need to know the mean and standard deviation before we can do this. We will standardize the variable AGE and possibly others.

$$\text{ZVAR} = (\text{VAR} - \text{mean})/\text{stddev}$$

Calculate Differences for Paired Pulse Measurements

$$\text{DIFFBA} = \text{PULSE2} - \text{PULSE1}$$

$$\text{DIFFAB} = \text{PULSE1} - \text{PULSE2}$$