Assignment for 8 - 6 - 10. Homework due at the start of class.

Estimating the difference μ_1 - μ_2 of the population means of two groups.

Case of paired comparisons. Scores

(x_i, y_i) $i \le n$

are observed on n subjects selected with equal probability and with replacement from a population.

Example A: Subjects are sampled from a population of patients who might benefit from a new medication. Before (x) and after (y) measurements of platelet counts are taken for each patient. These are possible dependent measurements. That is, a given patient's count after medication may be statistically linked to their count before medication. Give a 95% z-CI for the population mean difference score $\boldsymbol{\mu_D} = \mu_{before} - \mu_{after}$. Do this by ignoring the before and after scores and simply work with the difference scores d = x-y directly: $\bar{d} \pm 1.96\ s_d\ /\ \sqrt{n}$ where n is the number of difference scores (same as the number of subjects).

before after difference

$$\begin{pmatrix} 1367 & 1839 & -473 \\ 723 & 946 & -224 \\ 217 & 274 & -58 \\ 310 & 395 & -85 \\ 39 & 53 & -14 \\ 1362 & 1832 & -471 \\ 210 & 265 & -56 \\ 1 & 10 & -9 \\ 932 & 1233 & -302 \\ 1237 & 1656 & -420 \\ 1059 & 1409 & -350 \\ 37 & 51 & -14 \\ 63 & 81 & -19 \\ 397 & 509 & -112 \\ 33 & 46 & -13 \\ 543 & 703 & -161 \\ 203 & 256 & -54 \\ 214 & 270 & -57 \\ 7 & 16 & -10 \\ 37 & 50 & -14 \\ 1231 & 1649 & -418 \\ 1379 & 1856 & -477 \\ 1478 & 1995 & -518 \\ 836 & 1102 & -266 \\ 1453 & 1960 & -508 \\ 105 & 132 & -28 \\ 561 & 728 & -167 \\ 127 & 160 & -34 \\ 1014 & 1347 & -333 \\ 177 & 224 & -47 \\ 84 & 107 & -23 \\ 98 & 124 & -27 \\ 1 & 10 & -9 \\ 329 & 420 & -91 \\ 276 & 351 & -75 \\ 14 & 24 & -10 \\ 1112 & 1482 & -371 \end{pmatrix}$$

Case of independent samples

independent samples x_i $i \le n_x$ from population of x

independent samples y_i $i \le n_y$ from population of y

with x samples all independent of y samples.

Example B. Related to example A above except that we sample a number of subjects to receive a placebo and, independently of these, another sample of subjects to receive the medication. If the population of potential subects is large compared with the numbers sampled then approximate independence is achieved.

placebo {1460, 117, 419, 364, 555, 66, 207, 575, 698, 94, 648, 669, 261, 676, 200, 1157, 8, 282, 17, 1069, 609, 0, 0, 1548, 700, 859, 398, 410, 8, 447, 31, 1314}

med {164, 136, 842, 339, 94, 726, 91, 768, 466, 26, 1943, 78, 920, 1886, 10, 569, 1404, 1880, 751, 214, 35, 102, 64, 10, 537, 23, 479, 664, 944, 834, 17, 189, 1930, 20, 1260, 13}

It is natural to extimate the difference μ_x-μ_y of population means by the difference $\overline{x}$-$\overline{y}$ of sample means.

The standard deviation of $\overline{x}$-$\overline{y}$ is $\sqrt{\sigma_x^2 / n_x + \sigma_y^2 / n_y}$ because **$\overline{x}$ independent of $\overline{y}$** implies Var($\overline{x}$-$\overline{y}$) = Var($\overline{x}$) $\oplus$ Var($\overline{y}$). So a 95% z-CI is obtained by **plugging in estimates** in place of the unknwn population sigmas:

$$\overline{x} - \overline{y} \pm 1.96 \sqrt{s_x^2 / n_x + s_y^2 / n_y}$$

Homework due at the start of class 8-6-10.

1. Solve for the 95% z-CI for the difference μ_x-μ_y in the paried comparison of example A.

2. Solve for the 95% z-CI for the difference μ_x-μ_y in the independent samples setup of example B.

3. A table of data blocks A is posted separately for this assignment. Repeat (1) for your data block

A corresponding to your No. on the class list.

4. A table of data blocks B is posted separately forthis assignment. Repeat (2) for your data block B corresponding to your No. on the class list.