Test *t* za odvisna vzorca – metoda razlik

- odvisna vzorca oziroma N dvojic podatkov $\{X_1, X_2\}$ zberemo na različne načine, npr.
  - iste enote (npr. paciente ali laboratorijske preparate) merimo dvakrat (npr. pred in po zdravljenju oz. pred in po določenem postopku)
  - dvojčka, sorojenca, zakonca; leva in desna roka, sprednja in zadnja noga poskusne živali
  - izenačeni pari (npr. pacientov po spolu, starosti in stadiju bolezni)

- namesto, da bi analizirali opažene vrednosti, za vsako enoto izračunamo *razliko* med vrednostima spremenljivke (d_i=$X_{1i} - X_{2i}$; i=1..N)

  - opazimo, da je aritmetična sredina razlik enaka razliki aritmetičnih sredin:

$$
\begin{aligned}
M_1 - M_2 &= (\Sigma_{i=1..N}\, x_{1i} / N) - (\Sigma_{i=1..N}\, x_{2i} / N) \\
&= [(x_{11} + x_{12} + \ldots + x_{1N}) - (x_{21} + x_{22} + \ldots + x_{2N})] / N \\
&= (x_{11} - x_{21} + x_{12} - x_{22} + \ldots + x_{1N} - x_{2N}) / N \\
&= [\Sigma_{i=1..N}\, (x_{1i} - x_{2i})] / N \\
&= \Sigma_{i=1..N}\, d_i / N \\
&= M_d
\end{aligned}
$$

  - problem smo prevedli na preizkus razlike med opaženo aritmetično sredino in vnaprej določeno konstanto (test *t* za en vzorec)

  - izračunamo aritmetično sredino razlik (M_d); na podlagi standardnega odklona razlik (s_d) in velikosti vzorca ocenimo standardno napako ocene aritmetične sredine razlik: $SE_{M_d} = s_d / \sqrt{N}$

  - testna statistika je $t = (M_d - \mu_d) / SE_{M_d}$ (porazdeljuje se kot t z N–1 stopinjami prostosti)

  - če preizkušamo ničelno hipotezo, da med aritmetičnima sredinama ni razlike, t.j., da je populacijska aritmetična sredina razlik enaka nič (H0:μ_d=0), je testna statistika $t = M_d / SE_{M_d}$

- primer: dvanajst prostovoljcev je sodelovalo v študiji učinkovitosti diete, ki naj bi zmanjšala nivo holesterola; za vsakega je zabeležena koncentracija holesterola v krvnem serumu (v mg/dl) pred in po dieti; zanima nas, ali podatki dovolj trdno dokazujejo, da je dieta učinkovita?

  - podatki in izračun opisnih statistik:

prostovoljec	SH pred dieto (X_1)	SH po dieti (X_2)	razlika (d_i)
1	201	200	1
2	231	236	–5
3	221	216	5
4	260	233	27
5	228	224	4
6	237	216	21
7	326	296	30
8	235	195	40
9	240	207	33
10	267	247	20
11	284	210	74
12	201	209	–8
N = 12	M_1 = 244,25	M_2 = 224,08	M_d = 20,17 ; s_d = 23,13

- ničelna hipoteza: povprečna sprememba nivoja holesterola v serumu je enaka nič (H0:μ_d=0)

- testna statistika in vrednost p za dvosmerno testiranje:

  $t = 20{,}17 / (23{,}13 / 11^{1/2}) = 3{,}020$; p (df=11) = 0,012

- sklep: ničelno hipotezo zavrnemo na nivoju tveganja 5% (po dieti se povprečni nivo holesterola v serumu statistično značilno spremeni; ocenjujemo, da se v povprečju zmanjša za okoli 20 mg/dl)

- ocenimo 95% interval zaupanja za oceno povprečne razlike:

  $\mathrm{IZ}_{M_d} = M_d \pm t_{1-\alpha/2}\,\mathrm{SE}_{M_d} = 20{,}17 \pm (2{,}201 \times 6{,}68) = [5{,}47 ; 34{,}86]$

- namesto z metodo razlik, lahko razliko aritmetičnih sredin dveh odvisnih vzorcev analiziramo tudi z običajnim testom t (pri čemer imamo le N–1 stopinj prostosti)

  - pri oceni standardne napake ocene razlike aritmetičnih sredin moramo upoštevati kovarianco oziroma korelacijo med spremenljivkama: $t = (M_1 - M_2) / (\mathrm{SE}^2_{M_1} + \mathrm{SE}^2_{M_2} - 2\,r\,\mathrm{SE}_{M_1}\,\mathrm{SE}_{M_2})^{1/2}$

  - če bi uporabili test t za neodvisna vzorca, bi praviloma, t.j., če je korelacija med spremenljivkama pozitivna (kot npr. v zgornjem primeru, kjer je r=0,759), standardno napako ocene razlike aritmetičnih sredin precenili, torej bi imel test manjšo moč

- pri odločanju za načrtovanje študije, ki zahteva analizo odvisnih vzorcev, moramo glavno prednost tovrstne študije, t.j. izključitev čim večjega števila možnih zunanjih dejavnikov oz. nepreučevanih lastnosti, po katerih bi se lahko razlikovala neodvisna vzorca, pretehtati v primerjavi s slabostmi tovrstnega pristopa:

  - pri poskusih tipa pred-po, pri katerih ni možna randomizacija zaporedja pogojev, preučevanega učinka ne moremo ločiti od učinka učenja (oz. utrujenosti ipd.) in/ali učinka časa

  - iskanje ekvivalentnih parov (npr. pacientov) lahko zahteva veliko časa oz. sredstev

  - izguba stopinj prostosti (N-1 za N dvojic v primerjavi z 2N–2 za dva neodvisna vzorca velikosti N enot) in s tem izguba moči testa (ki pa ima za protiutež zgoraj opisano povečanje moči zaradi natančnejše ocene imenovalca testne statistike)