

Cvičení 7

1. Hodgkin Wilcoxon: Porovnat, zda se liší počty T4 a T8 buněk u pacientů s Hodgkinovým lymfomem (snímky 18-20 přednáška 10)
2. Deprese_Mann_whitney: Máme srovnat, zda se liší hodnoty jedné proměnné u dvou různých diagnóz (snímky 21-24 přednáška 10)
3. Alkohol: Posouzení vlivu alkoholu na schopnost řídit vozidlo – jedná se o párový design, protože porovnáváme situaci před a po požití alkoholu (snímky 15 a 16 přednáška 11)

Poznámky: Věnujte zvýšenou pozornost snímku 3 z přednášky 10. Tabulka Vás navede, jak vybrat vhodný test. Pokud jsou splněny předpoklady pro parametrický test (t-test párový či dvouvýběrový), použijeme ho (parametrické testy jsou silnější než neparametrické – snadněji prokážou rozdíl, pokud v datech nějaký je). Při porušení některého předpokladu, vybereme neparametrickou obdobu potřebného testu.

Parametrické testy vychází z nějakého teoretického rozložení: předp. že testová statistika se za platnosti nulové hypotézy řídí nějakým známým rozložením. Proto kritické hodnoty jsou kvantily konkrétních známých rozložení.

Na neparametrické testy mohou být také kladeny nějaké obecné předpoklady, ale nejedná se o předpoklady na konkrétní rozložení dat. Kritické hodnoty tedy nelze najít jako kvantily rozložení a jsou tabelovány pro konkrétní test. V excelu tyto tabulky nejsou implementovány (aspoň o tom zatím nevím), proto v případě neparametrických testů počítáme v podstatě „ručně“ a musíme porovnávat hodnotu testové statistiky s kritickou hodnotou z tabulek.

Pro většinu testů pro kategoriální data (testy v kontingenčních tabulkách) lze použít funkci chitest. Nc-Nemarův test je speciálním případem, který v excelu doposud implementován nebyl (novější verze už ho možná umí).