

POMŮCKA PRO INTERPRETACI STUDIÍ

P-HODNOTA (p-Value)

Většinou potřebujeme statistická data nejen k tomu, abychom mohli stručně popsat jejich vlastnosti, ale především proto, abychom mohli vyslovit závěr ohledně nějaké své domněnky – hypotézy. Například si můžeme myslet, že se za posledních deset let změnila inteligence studentů VŠ, a chceme to ověřit. Rozdíl mezi současnými studenty a průměrnou inteligencí tehdejších studentů je naše hypotéza, které ve statistice říkáme alternativní hypotéza – alternativní vůči nulové hypotéze. Nulová hypotéza je, že významný rozdíl neexistuje. Alternativní hypotéza v testu je naše domněnka, kterou se snažíme prokázat, nulová je její doplněk.

Test nám ukáže, zda je pravděpodobné, že platí naše alternativní hypotéza. Pracuje s tzv. hladinou signifikance neboli alfa – pravděpodobností, že jsme chybně zamítli nulovou hypotézu a přijali alternativní, která ve skutečnosti neplatí. Hladina je stanovena předem a většinou se pracuje s hodnotou 0,05 nebo 0,01. Vyjde-li nám v testu p hodnota, tedy konkrétní pravděpodobnost chybného zamítnutí nulové hypotézy pro použítá data menší, než dopředu stanovená alfa, je náš výsledek signifikantní – statisticky významný. Čím je p hodnota menší, tím pravděpodobnější je, že náš výsledek není chybný.

Srovnáváme dvě tvrzení:

H_0 – nulová hypotéza – většinou obhajuje stávající stav věcí.

H_A – alternativní hypotéza jí odporuje.

Chyba 1. druhu: nulová hypotéza sice platí, ale my ji zamítáme. Rozhodnutí o zamítnutí H_0 je dáno hladinou významnosti testu, což je maximální přípustná pravděpodobnost chyby 1. druhu. P-hodnota (dosažená hladina významnosti/signifikance) představuje jeden z nejméně používaných nástrojů při statistických analýzách v medicíně a s její pomocí je možno rozhodnout o zamítnutí nulové hypotézy na zvolené hladině významnosti. Hladina významnosti je pravděpodobnost chyby, že zamítneme platnou nulovou hypotézu. Čím nižší vyjde p-hodnota, tím více jsme přesvědčeni, že nulová hypotéza není správná a je třeba ji zamítnout. P-value je tedy taková nejnižší možná hladina významnosti, při které lze ještě zamítnout nulovou hypotézu.

Čím menší je p, tím neudržitelnější je nulová hypotéza, proto ji v takovém případě zamítáme.

Hladina významnosti alfa je hodnota, kterou si má zvolit experimentátor před analýzou dat. Obvykle volená hladina 0,05 vychází z historických konvencí.

P-hodnota nám neříká nic o velikosti pozorovaného efektu, nelze ji tedy použít pro porovnávání, ale je čistě vodítkem udržitelnosti či neudržitelnosti nulové hypotézy. Kdykoliv je to možné, doporučujeme použít raději interval spolehlivosti při interpretaci odhadu velikosti zkoumaného efektu než p-hodnotu.

Zdroj: Hynek M, Zvárová J. Statistické metody v medicíně II. – p-hodnota. Actual Gyn. 2016;8:24-25

INTERVAL SPOLEHLIVOSTI (Confidence Interval / CI)

Rozsah hodnot, ve kterém pravděpodobně leží průměr populace; obvykle pravděpodobnost 95%.

RELATIVNÍ RIZIKO (Relative Risk / RR)

Výpočet relativního rizika srovnává pravděpodobnosti výskytu sledovaného jevu ve dvou různých skupinách relativním způsobem, tedy tak, že dává do poměru pravděpodobnost výskytu jevu v experimentální skupině P_1 , a pravděpodobnost výskytu jevu v kontrolní skupině P_0 . Výsledkem je pak číslo, které vyjadřuje, kolikrát je v experimentální skupině vyšší pravděpodobnost výskytu sledovaného jevu než ve skupině kontrolní. $RR = 1$ znamená, že expozice sledovaným faktorem neovlivňuje vznik onemocnění.

$RR > 1$ znamená, že expozice sledovaným faktorem zvyšuje riziko vzniku nemoci, jde tedy o rizikový faktor.

$RR < 1$ znamená, že expozice sledovaným faktorem snižuje riziko vzniku nemoci, a daný faktor je tedy protektivní.

POMĚR ŠANCÍ (Odds Ratio / OR)

Pojem šance je často používán i v běžné mluvě, kde takto formou poměru hodnotíme možnou pravděpodobnost nastání nějakého jevu. Například šance na výhru v zápase 1 : 4 (jedna ku čtyřem) nebo šance uhodnutí určitého výsledku 1 : 9. Šance 1 : 4 se vztahuje k jevu, který nastane s pravděpodobností $1/(1 + 4) = 1/5$, tedy 0,20; obdobně šance 1 : 9 odráží jev s pravděpodobností 0,10. Šanci jevu A značíme $O(A)$ (odds) a definujeme ji jako podíl pravděpodobnosti nastání jevu A k pravděpodobnosti jevu opačného, tedy nenastání jevu A . Šanci ve prospěch A tedy vyjádříme jednoduchým vztahem

$$O(A) = \frac{P(A)}{1 - P(A)}$$

Šance výskytu jevu ve dvou populacích nebo v jinak definovaných skupinách osob (kohortách) lze dávat do poměru, potom hovoříme o **oměru šancí** (odds ratio = OR). Ten může nabývat nejnižší hodnoty 0, v maximu nijak ohraničen není. Vyjadřuje jednoduše, kolikrát je větší šance výskytu sledovaného jevu (např. nemoci) ve skupině osob exponovaných studovaným faktorem než ve skupině kontrolní (neexponované studovaným faktorem).

Zdroj: Dušek L. et al., Analýza dat v neurologii XXXV. Ces Slov Neurol N 2012; 75/108(5):642-645.

DIAGNOSTICKÉ TESTY

TP („true positive“ = správně pozitivní) – kolik výsledků bylo skutečně pozitivních (tzn. kolik pacientů bylo správně diagnostikováno jako pacienti).

FP („false positive“ = falešně pozitivní) – kolik výsledků bylo chybně pozitivních (tzn. kolik zdravých jedinců bylo chybně diagnostikováno jako pacienti).

FN („false negative“ = falešně negativní) – kolik výsledků bylo chybně negativních (tzn. kolik pacientů bylo chybně diagnostikováno jako zdraví).

TN („true negative“ = správně negativní) – kolik výsledků bylo skutečně negativních (tzn. kolik zdravých lidí bylo správně diagnostikováno jako zdraví).

SENZITIVITA (sensitivity)

schopnost testu správně diagnostikovat pacienta jako nemocného člověka. Vyjadřuje pravděpodobnost, že test bude pozitivní u skutečně nemocných lidí. Počítáme ji jako podíl počtu pacientů správně pozitivních v testu mezi všemi nemocnými.

Senzitivita testu = $TP : (TP + FN)$

SPECIFICITA (specificity)

schopnost testu rozpoznat zdravé pacienty, tj. stanovit u nich negativní výsledek. Vyjadřuje pravděpodobnost, že test bude negativní u zdravých lidí. Počítáme ji jako podíl počtu správně negativních v testu mezi všemi zdravými.

Specificita testu = $TN : (FP + TN)$

Pokud obě veličiny nabývají vysokých hodnot, test je dobrý pro hodnocení přítomnosti onemocnění. Naopak, pokud obě veličiny nabývají nízkých hodnot, test není dobrý pro hodnocení přítomnosti nemoci.

AUROC = area under receiving operation curves

ROC křivka = nástroj pro hodnocení klasifikačního testu, který ukazuje vztah mezi specificitou a senzitivitou daného testu. Kvalita se hodnotí podle plochy pod křivkou, nabývá hodnot od 0 do 1: 0,75 – 0,92 dobrá; 0,92 – 0,97 velmi dobrá; 0,97 – 1,00 vynikající

VĚROHODNOSTNÍ POMĚR (Likelihood Ratio - LR)

Věrohodnostní poměr značený jako LR+ vyjadřuje schopnost pozitivního výsledku testu odlišit nemocného a zdravého jedince.

Naopak věrohodnostní poměr značený jako LR- vyjadřuje obrácenou situaci, kdy test mylně označí nemocného jedince za zdravého.

U kvalitního testu je LR+ žádoucí co nejvyšší a naopak LR- co nejnižší. Vysoká hodnota LR+ jasně ukazuje na test, který poskytuje vysokou jistotu v určení nemoci, i přesto však není 100% zárukou, že osoba s pozitivním testem musí být vždy nemocná.

Zdroj: Dušek L. et al., Analýza dat v neurologii XXV. Ces Slov Neurol N 2012; 75/107(1):67-103.

Připravila: Mgr. J. Potomková, Ph.D. 2019, jarmila.potomkova@upol.cz