

Teoretické ukotvení

Diabetes mellitus neboli cukrovka je název pro skupinu chronických autoimunitních onemocnění, které jsou charakteristické poruchou metabolismu sacharidů. Mezi faktory podporující vznik cukrovky přispívá například vyšší věk (Laakso et al., 1985). Cukrovka má také větší prevalenci u žen a s vyšším počtem těhotenství je vyšší riziko onemocnění (Morris, 2022), mezi další rizikový faktor patří vysoký krevní tlak (Lago et al., 2007). Činitelů figurujících při vzniku cukrovky je více, avšak tyto se považují za základní a postačí k analýze, kterou si představíme níže.

Způsob zpracování dat

Data budeme zpracovávat binomální logistickou regresí, která nachází své uplatnění při situaci, ve které zkoumáme vztah mezi závislou proměnnou a jednou či více nezávislými proměnnými. Závislá proměnná je alternativní, nabývá tedy pouze hodnot 1 a 0, v tomto případě je tedy **závislá proměnná** (ne)přítomnost onemocnění: **1 = trpí cukrovkou, 0 = netrpí cukrovkou** (Hosmer & Lemeshow, 2000). Pro úplnou přehlednost je ještě nutné uvést, že konkrétní datový set zkoumal pouze ženy.

Nezávislé proměnné použité v modelu vypadají následovně:

- **Těhotenství** (počet kolikrát byla respondentka těhotná)
- **Glukóza** (plazmatická koncentrace glukózy a 2 hodiny v orálním glukózovém tolerančním testu)
- **Krevní tlak**
- **Tloušťka kůže** (měřená na tricepsu, udávaná v mm)
- **Insulin**
- **Věk**

Datový soubor čítající 768 participantů byl podroben binomální logistické regresy s použitím proměnných dle výše uvedeného popisu. V následující tabulce (Tabulka č. 1) jsou vyobrazeny výsledky analýzy. Pro účely této analýzy byla použita Waldova statistika, v tabulce jsou uvedeny hodnoty regresivních koeficientů a p hodnoty.

Tabulka 1: Výsledky logistické regrese

Model	Regresní koeficient	Testová statistika	p hodnota
Těhotenství	0,896299	12,8761	0,000333
Glukóza	0,962659	110,9077	0,000000
Krevní tlak	1,007943	2,7825	0,095302
Tloušťka kůže	0,983717	6,5080	0,010739
Insulin	1,001150	1,7649	0,184019
Věk	0,987572	1,9050	0,167520

* šedou barvou jsou zvýrazněny signifikantní výsledky

Signifikantní výsledky jsou prokázány u 3 proměnných. Konkrétně u těhotenství, respektive počtu dětí respondentky; glukózy a také u tloušťky kůže. Všechny tyto proměnné mají nízkou p hodnotu a také Waldova statistika dosahuje vysokých hodnot. Na podkladu těchto dat můžeme označit, že tyto 3 proměnné mají významnou souvislost s výskytem cukrovky.

Ukazatelé kvality modelu

V rámci zjišťování kvality modelu jsou použity dva ukazatele, jde konkrétně o Cox-Snell R^2 a Nagelkerke R^2 . Hodnota Cox-Snell R^2 nabývá čísla 0,24, což vysvětluje přibližně 24 %, Nagelkerke R^2 je rovno 0,38, vysvětluje tedy 38 %. Celkově zmíněné dva ukazatele vypovídají o relativně dobré přizpůsobivosti modelu k datům a umožňuje nám tak poměrně přesnou možnost predikce.

Zdroje:

1. Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* / David W. Hosmer, Stanley Lemeshow.
2. Laakso, M., Pyörälä, K., Laakso, M., & Pyörälä, K. (1985). Age of onset and type of diabetes. *Diabetes Care*, 8(2), 114-117. <https://doi.org/10.2337/diacare.8.2.114>
3. Lago, R. M., Singh, P. P., & Nesto, R. W. (2007). Diabetes and hypertension. *Nature Clinical Practice Endocrinology*, 3(10), 667-667. <https://doi.org/10.1038/ncpendmet0638>
4. Morris, D. (2022). Early-onset type 2 diabetes: Clinical implications, diagnosis and management. *Journal of Diabetes Nursing*, 26(6), 1-9.

Data a další informace o této zprávě jsou dostupné na adrese <https://dostal.vyzkum-psychologie.cz/stat4?i=100>.