

1 LINEÁRNÍ STATISTICKÝ MODEL

1.1 Teoretické ukotvení

Jedním z kritických faktů v epidemiologii alkoholismu je snižující se věk při zahájení konzumace alkoholu. Unnikrishnan Reghukumaran Nair, K Vidhukumar a Anil Prabhakaran dělali roku 2016 vědecké bádání, kde se snažili zjistit, zda došlo v České republice k poklesu věku při nástupu užívání alkoholu, a to v průběhu posledních 60 let. Informace čerpaly z dat, jež hovořily o počtu lidí, které vyhledaly odvykací služby ve fakultní nemocnici. Konkrétně na základě získaných údajů od 700 náhodně vybraných lidí, kteří jsou v procesu odvykání. Studie byla provedena na státní lékařské fakultě. Hlavními zkoumanými nezávislými proměnnými byly: rok narození, rodinná anamnéza, poruchy způsobené užíváním alkoholu a psychiatrická komorbidita. Zmíněná data byla testována skrze lineární regresi. Výsledky bádání říkají, že byl zjištěn významný lineární pokles věku při nástupu užívání alkoholu a poruchy užívání. Průměrný věk při nástupu užívání alkoholu se snížil z 24 na 17 let, resp. ze 46 na 21 let od kohorty narozených před rokem 1950 do kohorty narozených po roce 1985. Tento zlom nastal v 60. letech 20. století (Nair, Vidhukumar, Prabhakaran, 2016).

1.2 Lineární regrese

U lineární regrese se zabýváme takovými modely, jež popisují chování jedné závislé náhodné veličiny Y . Zároveň u ní vycházíme z toho, že je ovlivňována jedním až několika faktory $X_1, X_2, X_3 \dots$ (Dostál, 2021). Já se rozhodla zkoumat, jaký je vztah mezi mírou závislosti na alkoholu (Y) a pohlavím (X_1) a věkem respondentů (X_2).

[V mém modelu se nachází:](#)

Závislá proměnná:

Výsledný hrubý skór, který měří míru závislosti na alkoholu.

Regresory:

Spojité proměnná: věk.

Nominální proměnná: pohlaví (muž 1, žena 0).

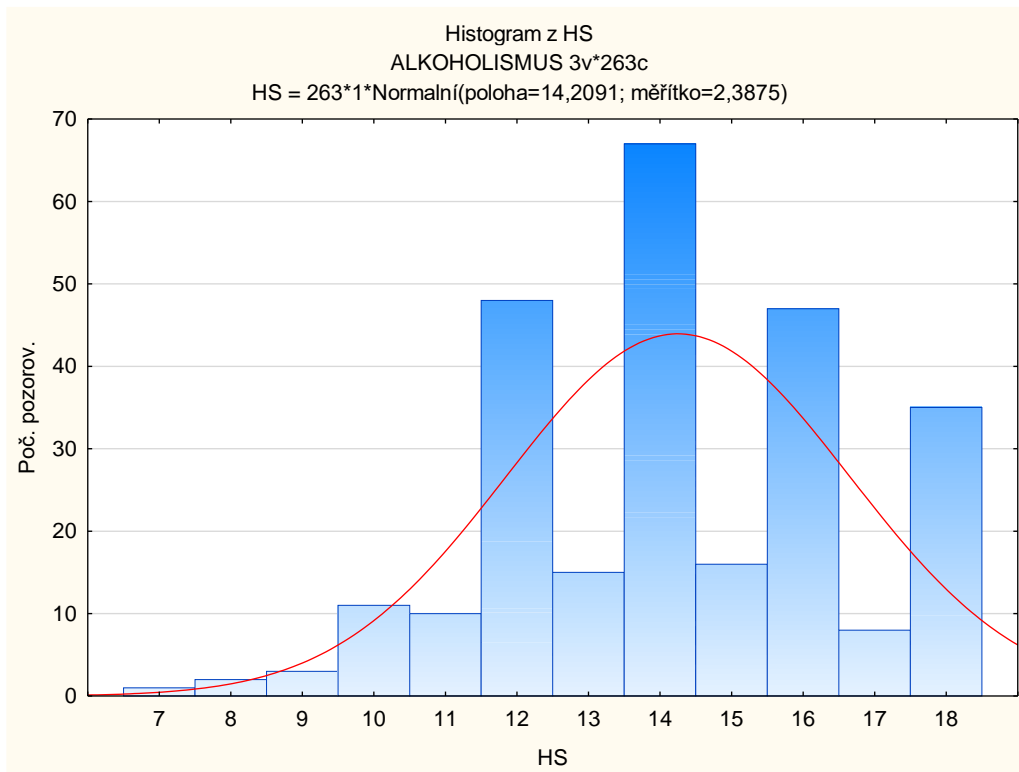
Vlastnosti výzkumného souboru:

Celkově máme údaje od 260 respondentů ve věkovém rozmezí 19 až 69 let. Konkrétně se v souboru nachází 204 žen a 56 mužů. Dané údaje jsem čerpala z výzkumného projektu, kde studenti tvořili vlastní testy v rámci předmětu Psychometrika. Konkrétně se jedná o test Alkoholismu od Venduly Cígerové.

Tabulku s daty najdete na tomto odkaze: <https://dostal.vyzkum-psychoologie.cz/stat4/data/data78.xlsx>

K užití lineární regrese je nutné splnit hned několik podmínek. Jedna z nich je fakt, že hrubý skór má normální rozdělení. Níže přikládám obrázek, jež skrze histogram znázorňuje normální rozložení hrubého skóru našeho modelu.

Obrázek č. 1: *Histogram HS závislosti na alkoholu*



1.3 Výsledky

K analýze dat jsem použila software TIBCO Statistica, verze 14.

Tabulka č. 1:

N = 260	b*	b	t (260)	p- hodnota
POHLAVÍ	- 0,111784	-0,64249	-1,82774	0,068734
VĚK	0,122096	0,02688	1,99635	0,046939

Tabulka č. 2:

R	R ²	F (2, 260)
0,16575370	0,02747429	3, 6726

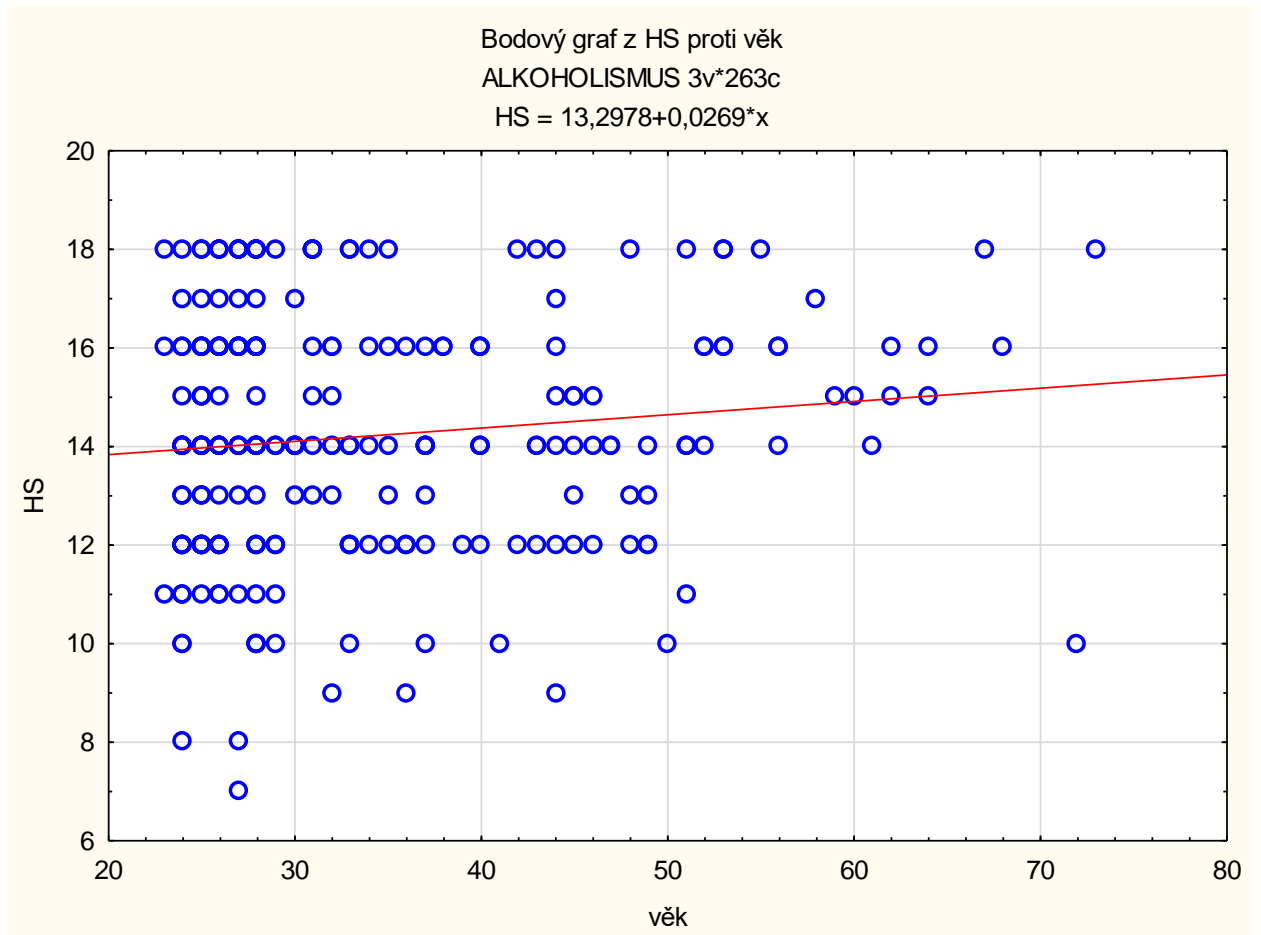
Z tabulek dokážeme vyčíst hned několik ukazatelů:

- **R² – koeficient determinace**, ten říká, kolik závislé proměnné dokážeme vysvětlit (Dostál, 2021). V našem případě dokážeme vysvětlit pouze 0,01999332. Takto malá hodnota značí, že náš model nedokáže vysvětlit téměř žádný rozptyl. Tento fakt si vysvětluji tím, že jednak máme v našem souboru málo respondentů, rovněž se zabýváme malým množstvím potřebných faktorů, domnívám se, že abychom popsali lépe danou problematiku, měli bychom se zabývat více faktory. Také neméně důležitý vliv může mít kvalita testu, .
- **b – nestandardizovaný regresní koeficient** říká o kolik se v průměru změní velikost závisle proměnné Y, když hodnota příslušného regresoru vzroste o jedničku (Dostál, 2021). Náš model tedy říká, že o 0,0268 bodů se zvedne hodnota hrubé skóru, když hodnota věku vzroste o jedničku.
- **b* standardizovaný regresní koeficient** získáme tak, že závislou proměnnou a regresory převedeme do formy z-skóru. Standardizovaný regresní koeficient říká, o kolik směrodatných odchylek v průměru vzroste hodnota závisle proměnné Y, pokud hodnota příslušného regresoru vzroste o jednu směrodatnou odchylku. Mj. se dají rovněž využít k porovnání efektu u jednotlivých regresorů (Dostál, 2021). Z výsledků našeho modelu je zřejmé, že když se posunu o jednu směrodatnou odchylku ve věkové kategorii, pak se výsledek testu posune asi o jednu desetinu směrodatné odchylky. Rovněž je zřejmé, že věk má o něco vyšší vliv než pohlaví.
- **Testová statistika F** ověřuje platnost hypotézy. Říká, že mezi predikcí a tou skutečnou hodnotou (skutečným stupněm míry závislosti) je souvislost, tzn., že nekorelují náhodně a náš model je tak schopen něco predikovat. V našem souboru vyšla $F(2, 260) = 3,6726$, s p hodnotou vyšší jak 0,0267.

1.4 Závěr:

Na závěr demonstruji výsledky skrze bodový graf hrubého skóru, jež hovoří o míře závislosti na alkoholu a věku respondentů.

Obrázek č. 2: Vztah hrubého skóru a věku respondentů



Zdroje informací:

- Dostál, D. (2016). *Lineární statistické metody v Psychologii*. Olomouc: Univerzita Palackého v Olomouci. Získáno z: https://dostal.vyzkum-psychologie.cz/soubory/skripta_linearni_modely.pdf
- Unnikrishnan Reghukumaran Nair, K Vidhukumar, & Anil Prabhakaran. (2016). Age at onset of alcohol use and alcohol use disorder: Time-trend study in patients seeking de-addiction services in Kerala. *Indian Journal of Psychological Medicine*, 38(4), 315–319. <https://doi.org/10.4103/0253-7176.185958>
- Cígenerová V., Test alkoholismu. Projekt v rámci předmětu Psychometrika 1.