

STANOVENÍ CUT-OFF SKÓRE U INVENTÁŘE SKRYTÝCH KUŘÁKŮ¹

Dokážeme podle chování rozpoznat, kdo je kuřák a kdo ne? Výzkumy naznačují, že kouření, ať už klasických cigaret, nebo jiných tabákových či nikotinových výrobků, může souviset s určitými osobnostními rysy a behaviorálními projevy, které kuřáky odlišují od nekuřáků. U kuřáků se například častěji objevují paranoidní, histriónské a antisociální rysy osobnosti (Becoňa et al., 2013). Podle výzkumů bývají kuřáci také častěji extravertnější, úzkostnější a impulzivnější než nekuřáci (Rondina Rde et al., 2007). Předpokládá se rovněž, že kuřáci při rozhovoru gestikulují dominantní rukou podobně, jako když v ní drží cigaretu (Gritz a Jarvik, 1978).

V rámci předmětu Psychometrika 1 vytvořili studenti **Inventář skrytých kuřáků** (Cmuntová et al., 2025). Cílem této zprávy bylo nalézt vhodné hraniční skóre (cut-off), na jehož základě bude možné určit, zda je u daného jedince pravděpodobná příslušnost ke skupině kuřáků či nikoli. Toto cut-off skóre bylo stanoveno pomocí **ROC analýzy**.

Pro analýzu byla využita data od 251 respondentů, kteří vyplnili výše uvedený inventář. Věk respondentů se pohyboval v rozmezí od 18 do 71 let, přičemž průměrný věk byl 29 let a medián 23 let. Výzkumný soubor tvořilo 187 mužů (75 %) a 64 žen (25 %).

Pro provedení ROC analýzy byly důležité následující **proměnné**:

- **informace**, zda je respondent v reálném životě kuřák (1) nebo nekuřák (0),
- **prevalence** kuřáků ve výzkumném souboru, která činila 20 %, a prevalence nekuřáků, která dosahovala 80 %,
- **hrubé skóre** v Inventáři skrytých kuřáků, přičemž minimální dosažitelné skóre bylo 12 bodů a maximální 48 bodů.

¹ Data a další informace o této zprávě jsou dostupné na adrese <https://dostal.vyzkum-psychologie.cz/stat4?i=634>.

Pro určení nejvhodnějšího cut-off skóre bylo nutné porovnat všechny možné hodnoty hrubého skóru, kterých respondenti při vyplňování Inventáře skrytých kuřáků skutečně dosáhli. Konkrétně se jednalo o hodnoty v rozmezí od 18 do 40 bodů. Na základě těchto hodnot bylo možné vypočítat několik podstatných ukazatelů, které sloužily k určení nejvhodnější hranice cut-off skóre.

Nejprve jsem stanovila počet jedinců spadajících do kategorie **TP (true positive)**, což znamená, že byli na základě zvoleného cut-off skóre správně klasifikováni jako kuřáci, a i ve skutečnosti kuřáky jsou. Dále bylo nutné určit počet jedinců označených podle cut-off jako kuřáci, kteří ve skutečnosti kuřáky nejsou. Tento případ označujeme jako **FP (false positive)**. Opačnou situaci představují jedinci, kteří byli podle cut-off skóre klasifikováni jako nekuřáci, ačkoli ve skutečnosti kuřáci jsou; tento případ označujeme jako **FN (false negative)**. Poslední kategorií je **TN (true negative)**, která vyjadřuje počet jedinců, které cut-off skóre správně zařadilo mezi nekuřáky, a kteří nekuřáky skutečně jsou.

Dále bylo zapotřebí vypočítat **senzitivitu** a **specificitu**. **Senzitivita** představuje poměr TP k celkovému počtu kuřáků a vyjadřuje pravděpodobnost, že zvolené cut-off skóre správně identifikuje kuřáka. **Specificita** je naopak poměr TN k celkovému počtu nekuřáků a vyjadřuje pravděpodobnost, s jakou cut-off skóre správně označí nekuřáka.

Na základě hodnot senzitivity a specificity lze vypočítat také **Youdenovu statistiku (J)**, přičemž vyšší hodnota tohoto ukazatele značí vhodnější volbu cut-off skóre. Youdenova statistika je však nejvhodnější v případě, kdy jsou skupiny rovnoměrně zastoupeny. Vzhledem k tomu, že ve výzkumném souboru výrazně převažují nekuřáci nad kuřáky, zařadila jsem proto navíc **statistiku I**, která zohledňuje poměr kuřáků a nekuřáků. Interpretace této statistiky je obdobná jako u Youdenovy statistiky – vyšší kladná hodnota znamená vhodnější cut-off skóre. Výsledky výše popsaných ukazatelů shrnuje **tabulka 1** níže.

Tabulka 1: Cut-off skóre

Cut-off	TP	FP	TN	FN	Senzitivita	Specificita	J	I
18	49	202	0	0	100 %	0 %	0,00	0,20
19	48	201	1	1	98 %	0 %	-0,02	0,20
20	48	200	2	1	98 %	1 %	-0,01	0,20
21	47	196	6	2	96 %	3 %	-0,01	0,21
22	47	193	9	2	96 %	4 %	0,00	0,22
23	47	189	13	2	96 %	6 %	0,02	0,24
24	44	177	25	5	90 %	12 %	0,02	0,27
25	43	161	41	6	88 %	20 %	0,08	0,33
26	40	142	60	9	82 %	30 %	0,11	0,40
27	35	123	79	14	71 %	39 %	0,11	0,45
28	34	105	97	15	69 %	48 %	0,17	0,52
29	34	90	112	15	69 %	55 %	0,25	0,58
30	27	72	130	22	55 %	64 %	0,19	0,63
31	21	55	147	28	43 %	73 %	0,16	0,67
32	16	39	163	33	33 %	81 %	0,13	0,71
33	13	24	178	36	27 %	88 %	0,15	0,76
34	10	14	188	39	20 %	93 %	0,13	0,79
35	8	11	191	41	16 %	95 %	0,11	0,79
36	6	5	197	43	12 %	98 %	0,10	0,81
37	6	3	199	43	12 %	99 %	0,11	0,82
38	3	1	201	46	6 %	100 %	0,06	0,81
39	1	1	201	48	2 %	100 %	0,02	0,80
40	0	1	201	49	0 %	100 %	0,00	0,80
41	0	0	202	49	0 %	100 %	0,00	0,80

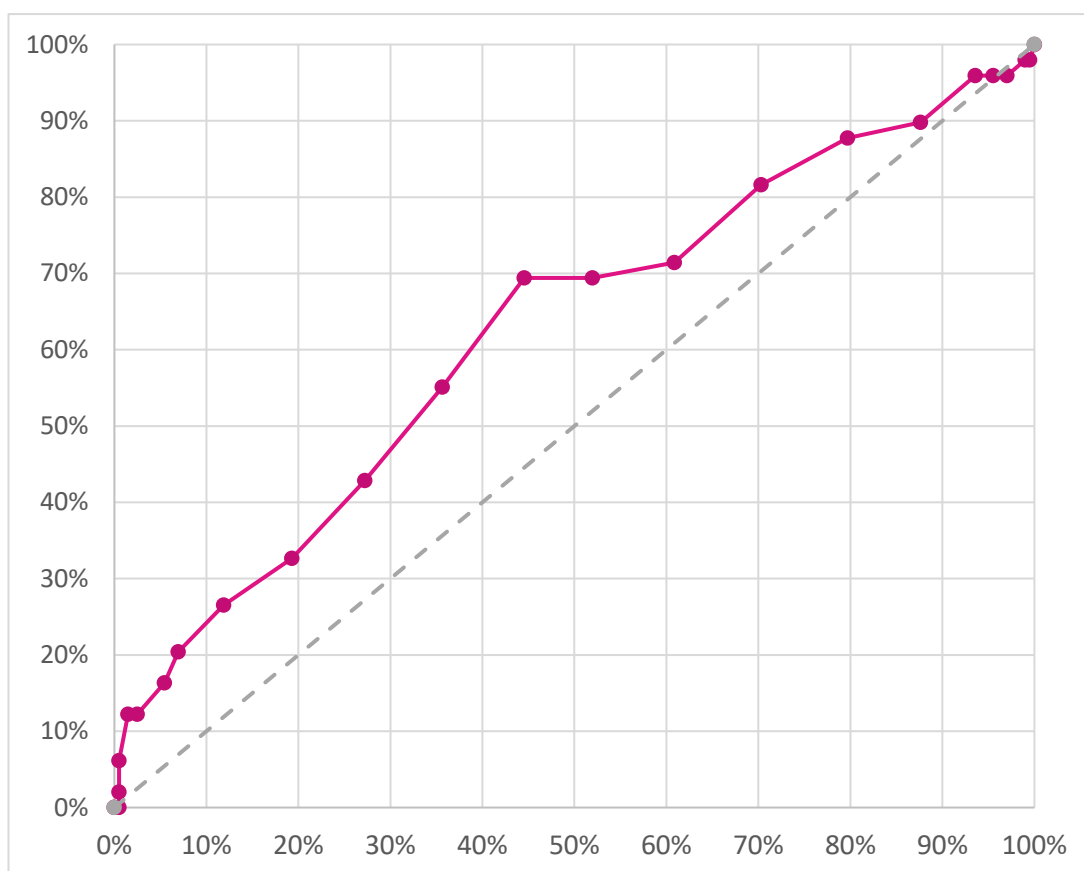
Nejvyšších kladných hodnot Youdenovy statistiky bylo dosaženo při **cut-off skóre 29**.

Z hlediska interpretace je však pro nás podstatnější hodnota statistiky I, která zohledňuje také prevalenci kuřáků a nekuřáků v souboru. Ta dosáhla maxima při **cut-off skóre 37**. Lze tedy předpokládat, že pokud jedinec v inventáři získá 37 a více bodů, je pravděpodobné, že patří do skupiny kuřáků.

V grafu 1 je znázorněna **ROC křivka**. V ideálním případě by se křivka měla stáčet co nejblíže k levému hornímu rohu grafu. Pokud by ROC křivka probíhala podél šedé diagonální linie, znamenalo by to, že test nerozlišuje mezi skupinami lépe než náhodná klasifikace (například hodem mincí). V případě Inventáře skrytých kuřáků je křivka mírně vychýlena nad tuto diagonálu, což naznačuje, že test má určitou diskriminační schopnost a že kuřáci v průměru dosahovali vyššího hrubého skóre než nekuřáci.

Plochu pod křivkou označujeme jako **AUC (area under the curve)**. V našem případě dosahuje hodnota **AUC = 0,559**. Tato hodnota naznačuje, že model má pouze omezenou schopnost rozlišovat mezi kuřáky a nekuřáky. Jelikož se hodnota AUC pohybuje jen mírně nad hranicí 0,5, která odpovídá náhodnému rozhodování, lze diskriminační schopnost modelu hodnotit jako slabou, což je v souladu se závěrečným hodnocením nástroje, který ve své zprávě prezentují i sami autoři. Podle autorů inventáře současná verze nesplňuje základní psychometrické předpoklady ani standardy kladené na psychologické screeningové metody.

Graf 1: ROC křivka



LITERATURA

Becoña, E., Fernández del Río, E., López-Durán, A., & Martínez, Ú. (2013). Do Cigarette Smokers Have Different Personality Patterns than Non-Smokers? *The Spanish Journal of Psychology*. <https://doi.org/10.1017/sjp.2013.101>

Cmuntová, N., Kasíková, L., Linhartová, J., Volosová, V. A., & Moyzes, V. (2025). *INVENTÁŘ SKRYTÝCH KUŘÁKŮ*. PhDr. Daniel Dostál, Ph.D. https://dostal.vyzkum-psychologie.cz/pmlab/zpravy/zprava0324_4.pdf

Gritz, E.R., Jarvik, M.E. (1978). Nicotine and Smoking. In: Iversen, L.L., Iversen, S.D., Snyder, S.H. (eds) *Stimulants*. Handbook of Psychopharmacology, 11. Springer, Boston, MA. https://doi.org/10.1007/978-1-4757-0510-2_9

Rondina Rde, C., Gorayeb, R., & Botelho, C. (2007). Psychological characteristics associated with tobacco smoking behavior. *J Bras Pneumol*, 33(5). <https://doi.org/10.1590/s1806-37132007000500016>