8_Testování hypotéz

Testování hypotéz: filozofický úvod

Výrok: „Všechny vrány jsou černé“

Je možné dokázat tento výrok?

Nikoli - nelze pozorovat všechny vrány

Řešení = Falzifikace výroku: pokud najdeme nečernou vránu, ukážeme že ne všechny vrány jsou černé

Jediné pozorování může vést k závěru že výrok je nepravdivý: metoda *modus tollens*

H0: všechny vrány jsou černé
H1: ne všechny vrány jsou černé
Snažíme se falzifikovat H0

Aplikace na příklad s IQ

- Př. Výsledky IQ testu jsou aproximovány (blíží se) normálním rozložením o průměru $\mu = 100$ a $\sigma=16$. Třída 36 dětí dosáhne průměru 105 bodů. Učitel si myslí že děti patří do populace $\mu > 100$
- H 0: $\mu = 100$
  H 1: $\mu > 100$
- Snažíme se falzifikovat H0, statisticky řečeno zamítnout H0
- Kdy zamítáme? Když máme dostatečnou evidenci
- Zamítnutí H0 založíme na náhodě/riziku/šanci
  - Šanci, že najdeme výběrový průměr 105 nebo vyšší pokud je vzorek vybrán z populace s $\mu = 100$.
  - Pokud je šance malá (menší než α) zamítáme H0.

Účel testování hypotéz

- Pomocí pravděpodobnosti zjistit zda data ve vzorku podporují určitý předpoklad/stanovisko (=hypotézu) o populaci
- Statistická hypotéza = výrok o populaci (např. že parametr (průměr nebo proporce) nabývá určité konkrétní hodnoty)

Nulová a alternativní hypotéza

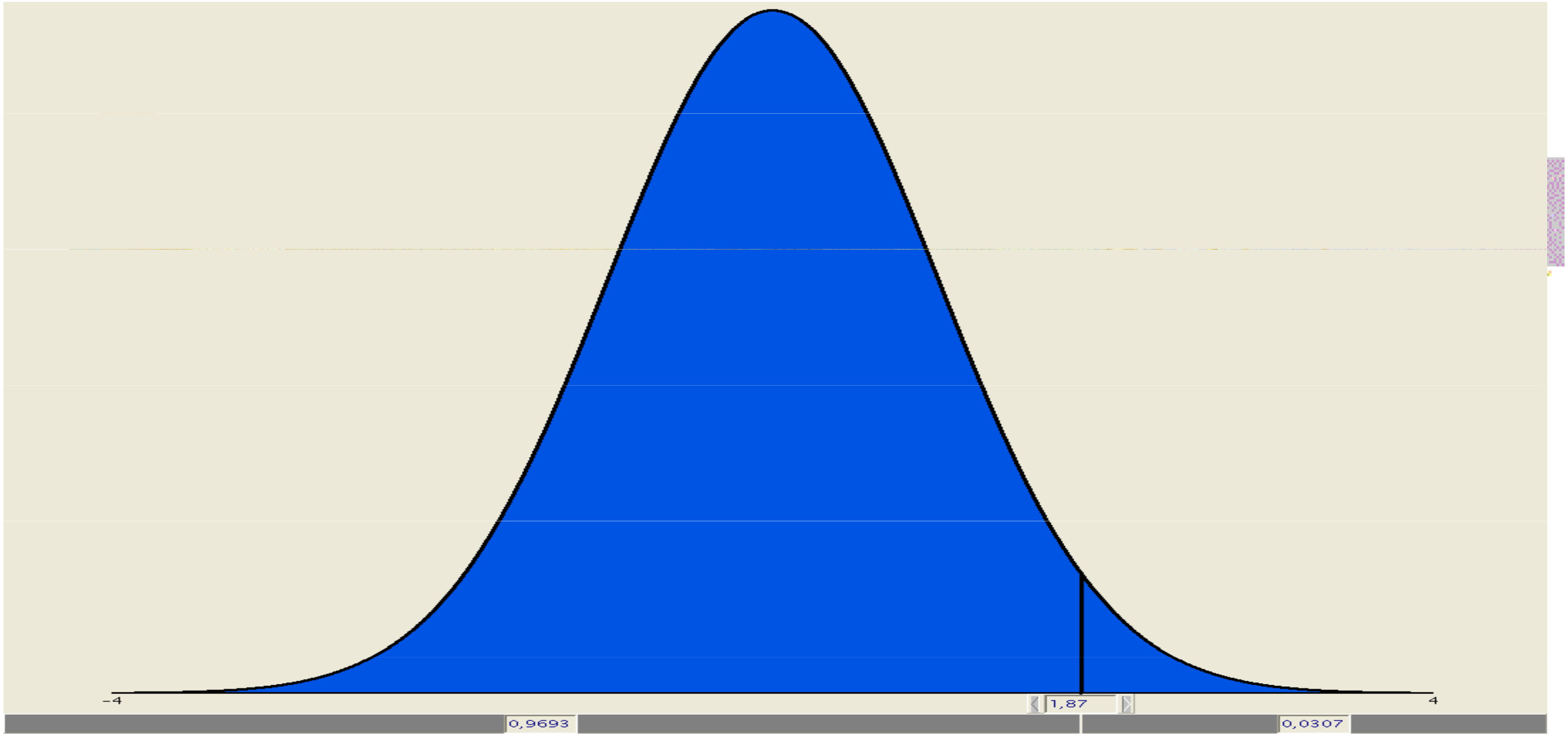
- Nulová hypotéza (H_0)
  - = tvrzení že parametr nabývá určité konkrétní hodnoty (např. parametry jsou si rovny, efekt proměnné je roven 0, mezi proměnnými v populaci není vztah)
  - Hodnota nulové hypotézy obvykle představuje absenci efektu
- Alternativní hypotéza (H_1)
  - = tvrzení, že parametr nabývá jiné hodnoty (obvykle nějakého intervalu hodnot) než v H_0
  - představuje efekt určitého druhu

Testovaná statistika (z nebo t-skór)

- = určuje jak daleko leží bodový odhad od populačního parametru za předpokladu že platí H0
- Vzdálenost měřena pomocí směrodatných chyb např. t=1,87 znamená že bodový odhad leží 1,87 SE od průměru
- Výpočet statistiky závisí na typu výzkumného problému (např. predikce hodnoty populační proporce nebo průměru)
  - H0: μ = 100, pak $$t = \frac{x - \mu_0}{s/\sqrt{n}}$$
  - H0: p = 0,26, pak $$z = \frac{(\hat{p} - p_0)}{se_0} = \frac{(\hat{p} - p_0)}{\sqrt{p_0(1 - p_0)/n}}$$

P - hodnota

- = podmíněná pravděpodobnost (pokud platí H0) že testovaná statistika (z nebo t) je rovna pozorované hodnotě nebo hodnotě extrémnější
- Čím je p-hodnota nižší, tím větší „důkaz“ proti H0 máme



Hladina významnosti (alfa-α)

- = určitá námi zvolená hodnota (pravděpodobnost / riziko)
- Pokud je p-hodnota testované statistiky nižší nebo rovna této hodnotě, pak zamítáme H_0 a říkáme , že výsledek je statisticky významný na hladině α
- Pokud je p-hodnota testované statistiky vyšší než tato hodnota, pak nezamítáme H_0
- Hladina významnosti se obvykle stanovuje na 0.05 nebo přísněji na 0.01

Testování hypotéz: postup

- 1. Učiníme předpoklad o hodnotě parametru, např. populačního průměru a tímto stanoví nulovou hypotézu H0
- 2. Za předpokladu že H0 je pravdivá, zkonstruujeme distribuci všech možných (potenciálních) hodnot statistiky vzorku, zde průměru vzorku, když vybírá jednoduchý náhodný vzorek o velikosti N z populace o průměru předpokládaném H0. Takto vzniká výběrová distribuce statistiky, zde výběrová distribuce průměru.Tvar této distribuce může být odvozen nezávisle na tvaru populační distribuce (centrální limitní věta) jako normální s průměrem rovným populačnímu průměru, přičemž tvar se blíži normálnímu s rostoucí velikostí vzorku
- 3. S výběrovou distribucí průměru stanovujeme podmíněnou pravděpodobnost (p-hodnotu), s jakou se daná nebo extrémnější hodnota výběrového průměru vyskytne za předpokladu že H0. je pravdivá
- 4. Pokud je p-hodnota nižší než α (alfa), tj. námi předem stanovená kritická hodnota (též riziko chyby 1.druhu), říkáme: „Pokud je H0 pravdivá, tak podmíněná pravděpodobnost že najdu danou hodnotu výběrového průměru nebo extrémnější je nižší než α. Tato pravděpodobnost je tak nízká (tj. je velmi neobvyklé, že bych dostal takovýto výsledek, pokud by H0 platila), že již nedůvěřujeme H0 a zamítáme ji.
- 5. Pokud je p-hodnota větší než α, pak říkáme: „Pokud je H0 pravdivá, tak podmíněná pravděpodobnost že najdu danou hodnotu výběrového průměru nebo extrémnější je vyšší než α. Proto nemáme potřebnou jistotu (dostatečnou evidenci) a H0 nezamítáme.

Formulace H1 a regiony zamítnutí

- Alternativní hypotézy H1 mohou být formulovány jednosměrně nebo obousměrně
- Př. Obousměrně : H0=100, H1≠100
  (průměry H1 mohou být menší nebo větší než 100, proto obousměrně)
- Př. Jednosměrně: H0=100,H1>100
  (zajímají nás průměry H1 pouze větší než 100, proto jednosměrně)

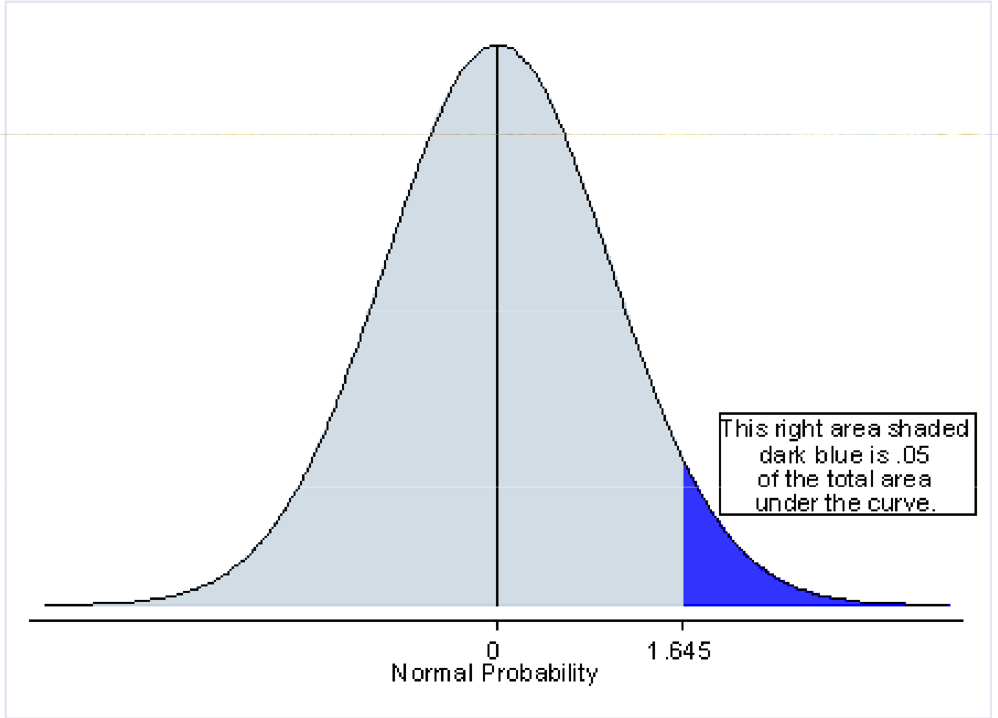
- Jednosměrně – jednostranný test α=0.05
- Obousměrně – oboustranný test α=0.05 = (0.025 spodní interval + 0.025 horní interval)

Tabulka z-hodnot pod 1-straným a 2-straným testem (viz tabulka z-skoru)

Test \ α	0.05	0.01	0.001
1-straný	1.645 F(z) = .95	2.33 F(z) = .99	3.09 F(z) = .999
2-straný	1.96 F(z) = .975	2.58 F(z) = .995	3.29 F(z) = .9995

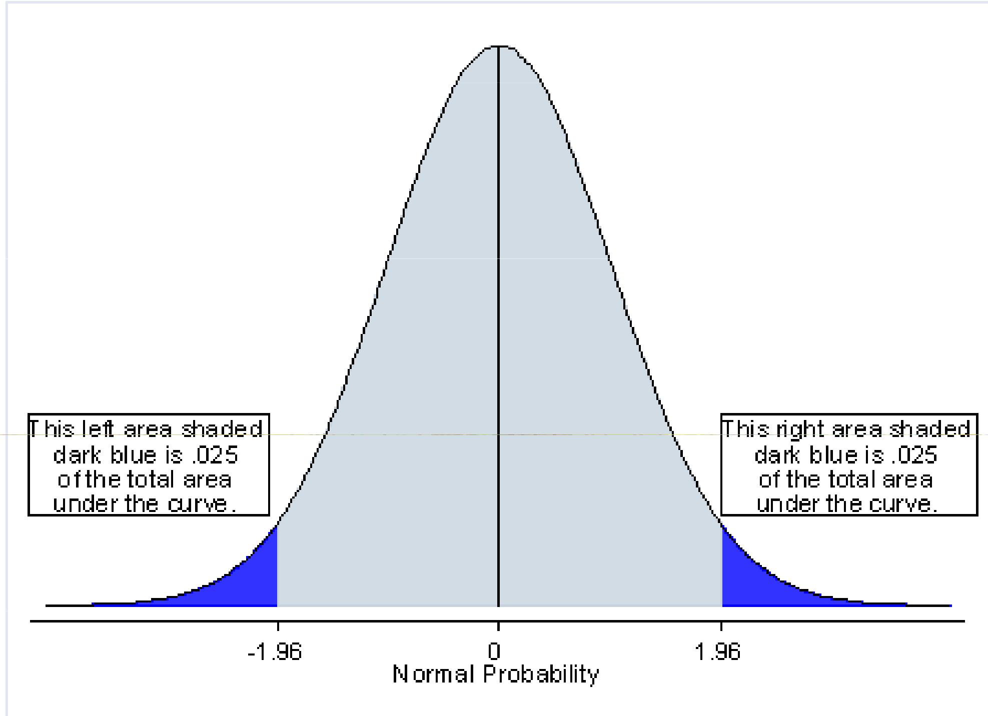
1-stranný test

•Jednosměrně – jednostranný test α=0.05



2-stranný test

oboustranný test pro α=0.05 = α / 2 = (0.025 spodní interval + 0.025 horní interval)



Chyby 1. a 2. druhu

- Ikdyž podmíněná pravděpodobnost vytáhnutí daného nebo extrémnějšího průměru vzorku bude menší než alfa a člověk důsledkem toho zamítne H_0, stále existuje pravděpodobnost, ikdyž nízká (menší než α), že daný průměr pochází z populace s H_0a člověk se tak rozhodnu špatně – učiní chybu 1. druhu.

- Stejně tak existuje situace kdy podmíněná pravděpodobnost vytáhnutí daného nebo extrémnějšího průměru vzorku bude větší než alfa a člověk důsledkem toho nezamítne H_0, ikdyž existuje pravděpodobnost (v závislosti na alternativní hypotéze H_1) že daný průměr pochází z populace H_1a člověk také učiní špatné rozhodnutí – chybu 2. druhu β.

Čtyři možné situace a síla testu

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="2">skutečnost</td></tr>
<tr><td>H0</td><td>H1</td></tr>
<tr><td rowspan="2">rozhodnutí</td><td>„H0“</td><td>p(„H0“ | H0)

1 – α</td><td>Chyba 2. typu (β)

riziko chybného nezamítnutí nulové hypotézy
p(„H0“ | H1)</td></tr>
<tr><td>„H1“</td><td>Chyba 1.typu (α)

riziko chybného zamítnutí nulové hypotézy
p(„H1“ | H0)</td><td>p(„H1“ | H1)

Síla testu (1 – β)</td></tr>
</table>