BIOSTATISTIKA

L05

Co nám statistický test dovoluje tvrdit?



Od velikosti efektu k testování hypotéz

Ilustrace vytvořená pomocí AI. Anatomická reference: Natalia Agudelo / SERC, CC BY 2.0.

cuni-natur-biostatistics.github.io/materialy.html

Měsíc je celý ze sýra!



Ilustrační obraz vytvořen AI.

Co by mohlo změnit váš závěr?

> Jaké pozorování by vás přimělo tvrzení o sýrovém Měsíci vyvrátit?



Která tvrzení mohou pozorování změnit?



Všechny labutě jsou bílé.



Holub stěhovavý vyhynul.



Měsíc je celý ze sýra.



Tento pták je nejkrásnější.

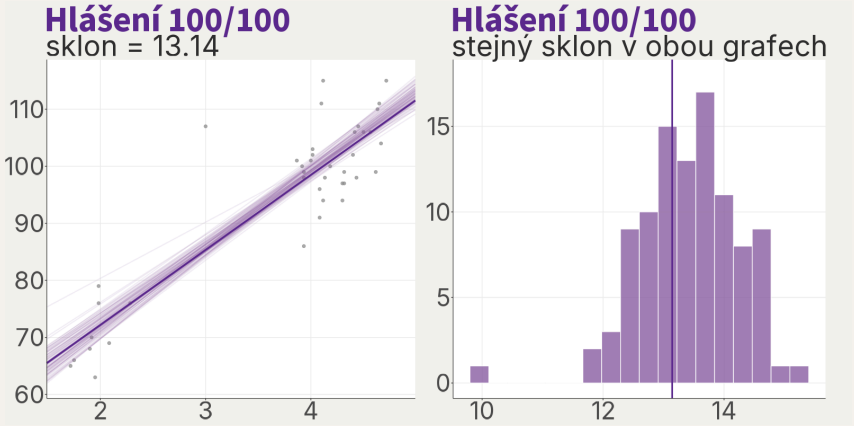
Která tvrzení může změnit nové pozorování? U kterého to nedává smysl?

Testovatelné neznamená pravděpodobné ani pravdivé. Pozorování jsou naše omezené informace o světě.

Co si pamatujete z minulé lekce?

Odpovězte samostatně na tři krátké otázky. Potom společně projdeme správné odpovědi.

Proč sto pozorovatelů nedostalo ze svých výběrů přesně stejný sklon?



- Skutečný vztah se po každém výpočtu změnil.
- Funkce lm() vybírá sklon náhodně.
- Každý pracoval s jiným výběrem pozorování.
- Sklon nemá s daty souvislost.

SPRÁVNÁ ODPOVĚĎ

: Každý pracoval s jiným výběrem pozorování.

Výběry obsahují různá pozorování, a proto se jejich odhady liší. Této změně mezi výběry říkáme výběrová variabilita.

Které číslo popisuje velikost odhadnutého vztahu mezi délkou erupce a čekáním?

Řádek sklonu z modelu

Parametr	Estimate	Std. Error
delka_erupce_min	13,34	0,66

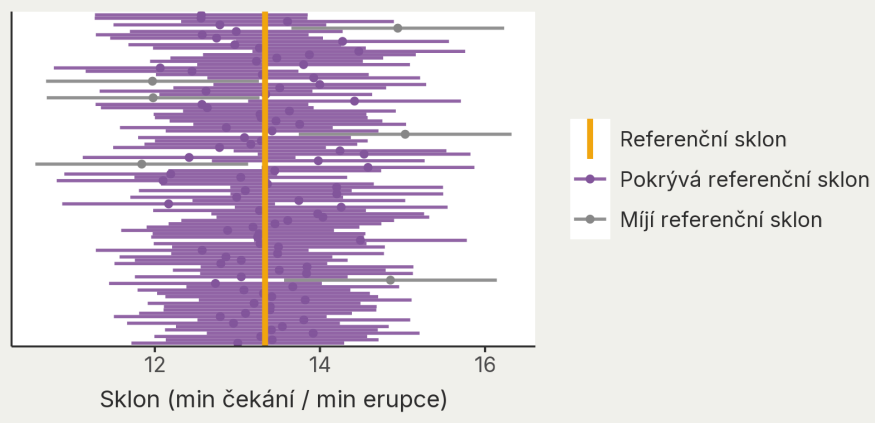
- Hodnota ve sloupci Std. Error
- Počet řádků tabulky
- Název prediktoru
- Hodnota ve sloupci Estimate

SPRÁVNÁ ODPOVĚĎ

: Hodnota ve sloupci Estimate

Estimate je odhad sklonu. Std. Error popisuje přesnost tohoto odhadu.

Co v tomto opakovaném postupu znamená označení „95% interval spolehlivosti“?



Při opakování stejného postupu přibližně 95 % takto vytvořených intervalů pokryje skutečný parametr.

Vypočtený interval má po výpočtu 95% pravděpodobnost obsahovat pevný parametr.

Uvnitř intervalu leží 95 % jednotlivých pozorování.

Odhad je s pravděpodobností 95 % bez chyby.

SPRÁVNÁ ODPOVĚĎ

: Při opakování stejného postupu přibližně 95 % takto vytvořených intervalů pokryje skutečný parametr.
95 % popisuje dlouhodobou úspěšnost postupu při opakovaných výběrech, nikoli pravděpodobnost jednoho již vypočteného intervalu.

Od intervalu k nulovému tvrzení

> Interval ukazuje hodnoty slučitelné s daty. **Jak slučitelná jsou data s jednou konkrétní nulovou hodnotou?**

Výsledky učení

Po absolvování této přednášky budete schopni:

- rozlišit biologickou realitu od toho, co zachytila konkrétní studie
- interpretovat pozorovaný efekt a jeho nejistotu
- převést biologické tvrzení na nulovou a alternativní hypotézu o sklonu
- vysvětlit, co vyjadřuje p-hodnota a co nevyjadřuje
- rozlišit chybu I. a II. typu a napsat opatrný vědecký závěr

Co z reality skutečně pozorujeme?

Seznamte se: krab houslista



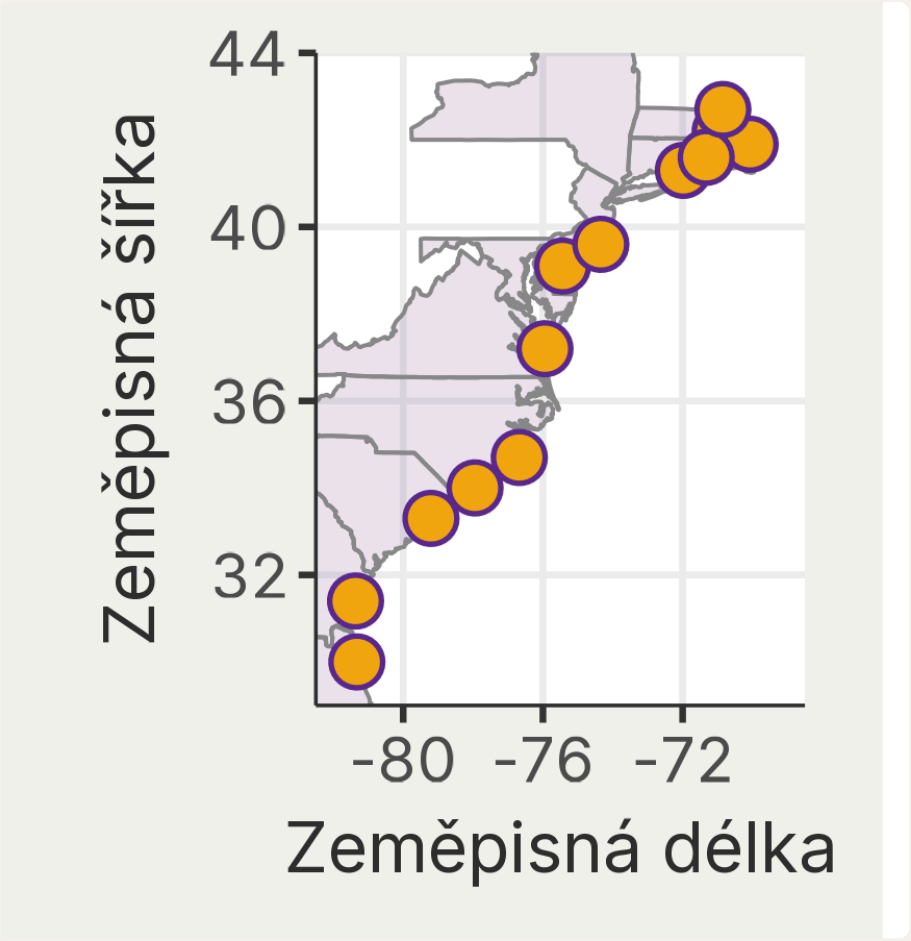
Minuca pugnax

Samec nese jedno nápadně zvětšené klepeto.

Žije ve slaniscích podél Atlantiku.

Foto: Natalia Agudelo / Smithsonian Environmental Research Center, Wikimedia Commons, CC BY 2.0.

Slaniska podél Atlantiku



CO BUDEME POZOROVAT

- 13 slanisek od Floridy po Massachusetts
- jeden bod = průměrná šířka krunýře na jedné lokalitě
- zeměpisná šířka a velikost v milimetrech

Jak se průměrná velikost mění směrem na sever?

Data: LTER `pie_crab`, CC0; souřadnice: Johnson et al. 2019.

Kterou proměnnou chceme vysvětlit?

ZEMĚPISNÁ ŠÍŘKA

Poloha slaniska ve stupních severní šířky

PRŮMĚRNÁ ŠÍŘKA KRUNÝŘE

Velikost samců na lokalitě v milimetrech

Která proměnná je **prediktor** a která **odezva**?

Bergmannovo pravidlo

Organismy bývají v chladnějších oblastech větší.

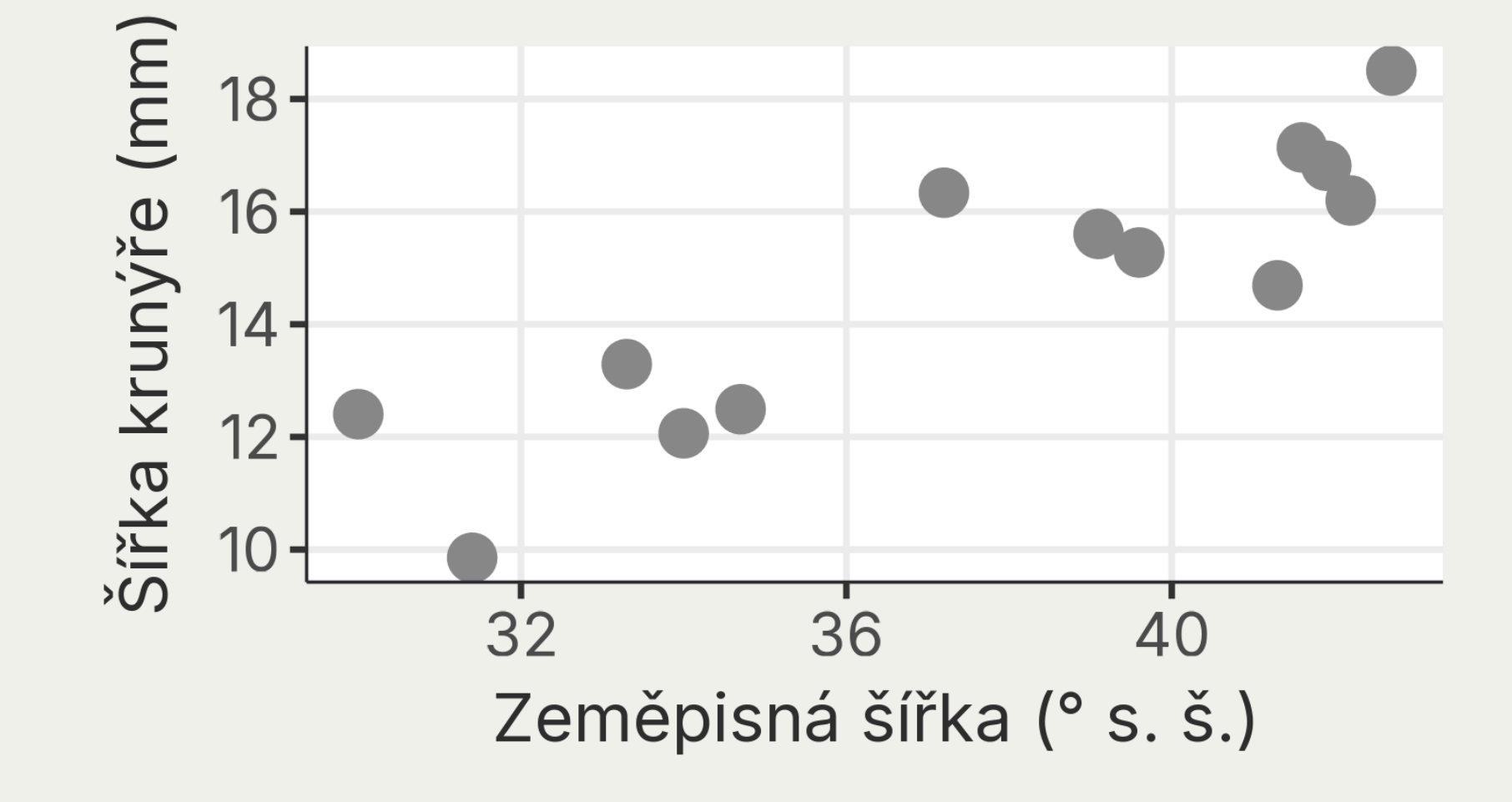
U krabů to nemusí platit. Je to biologická předpověď, kterou porovnáme s daty.

Pokud by byli krabi na severu větší, jaký směr vztahu bychom očekávali?

Klesající

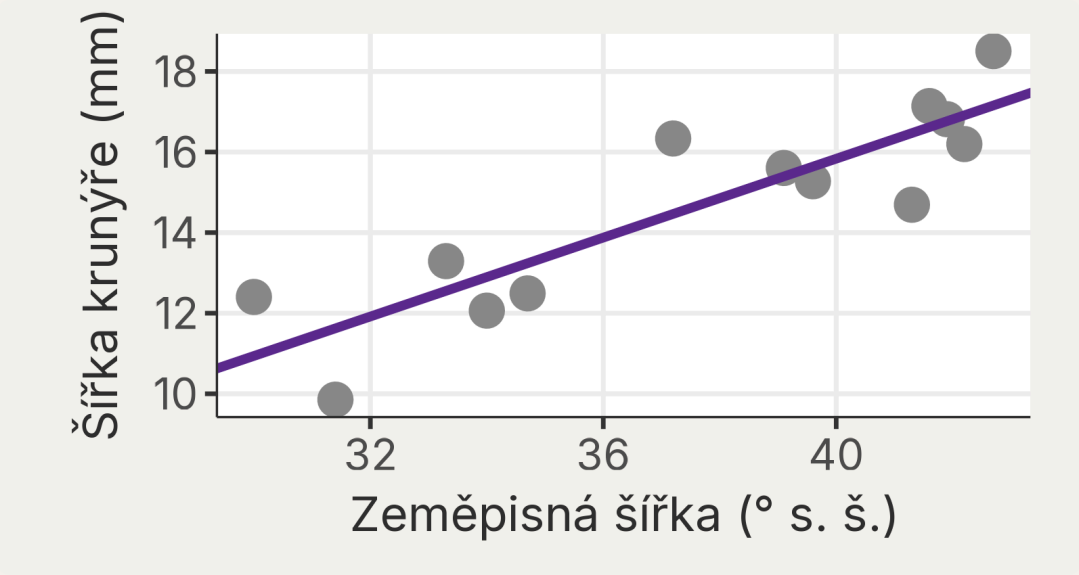
Rostoucí

Co ukazují data?



Popište směr a rozptyl bodů. Co je na obrazu přesvědčivé a co vás zneklidňuje?

Lineární model




```
mod_krabi <-
  lm(
    formula = prumerna_sirka_mm ~
      zemepisna_sirka,
    data = data_lokality
  )
```

summary()

```
1  summary(mod_krabi)
```

Parametr	Odhad	Std. chyba	t	p
Intercept	-3,747	3,009	-1,25	0,239
Zeměpisná šířka	0,490	0,079	6,16	<0,001

ODHADNUTÝ MODEL

Průměrná šířka = -3,747 + 0,490 × zeměpisná šířka + residuum

Z MINULÉ LEKCE

Estimate = intercept a sklon;
Std. Error = nejistota odhadu.

Co znamenají `t value` a `Pr(>|t|)`?

Je naměřený vztah realita?



Ilustrace vytvořená pomocí AI.

Je fialová přímka skutečný vztah v přírodě, nebo odhad z této studie?

Jak často by podobně silný vztah vznikl prostou náhodou ve světě bez systematického trendu?

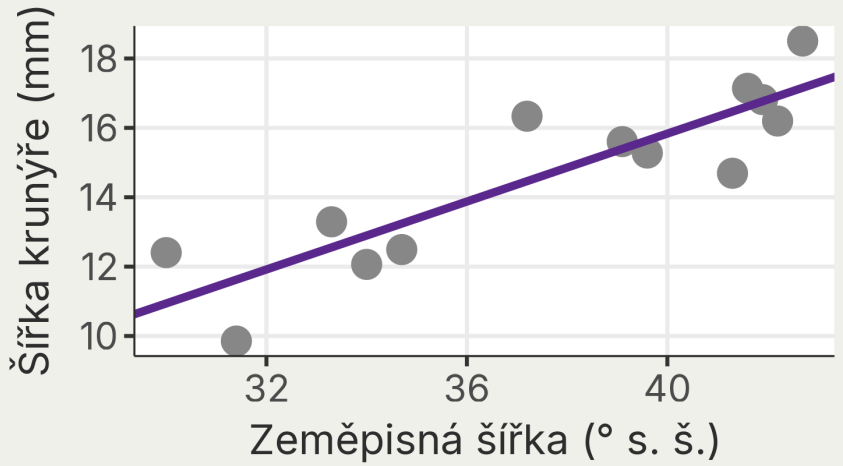
REALITA

?

Skutečný vztah ve všech možných lokalitách přímo nevidíme.

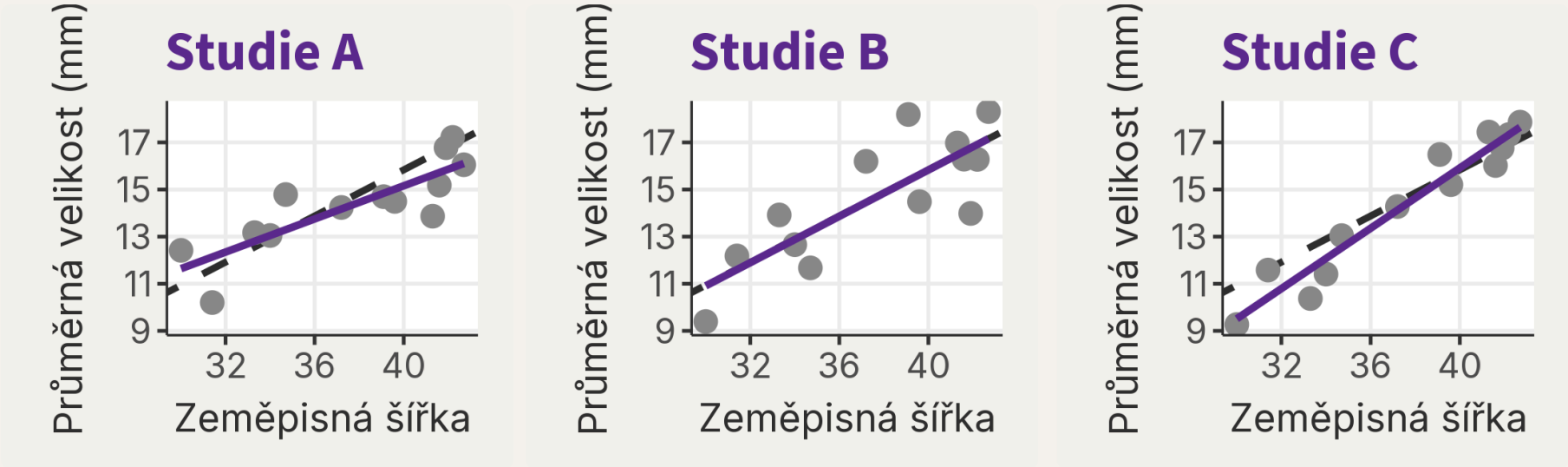
→

výběr, měření, místní rozdíly



NAŠE POZOROVÁNÍ: 13 lokalit

Stejná realita, jiné pozorování

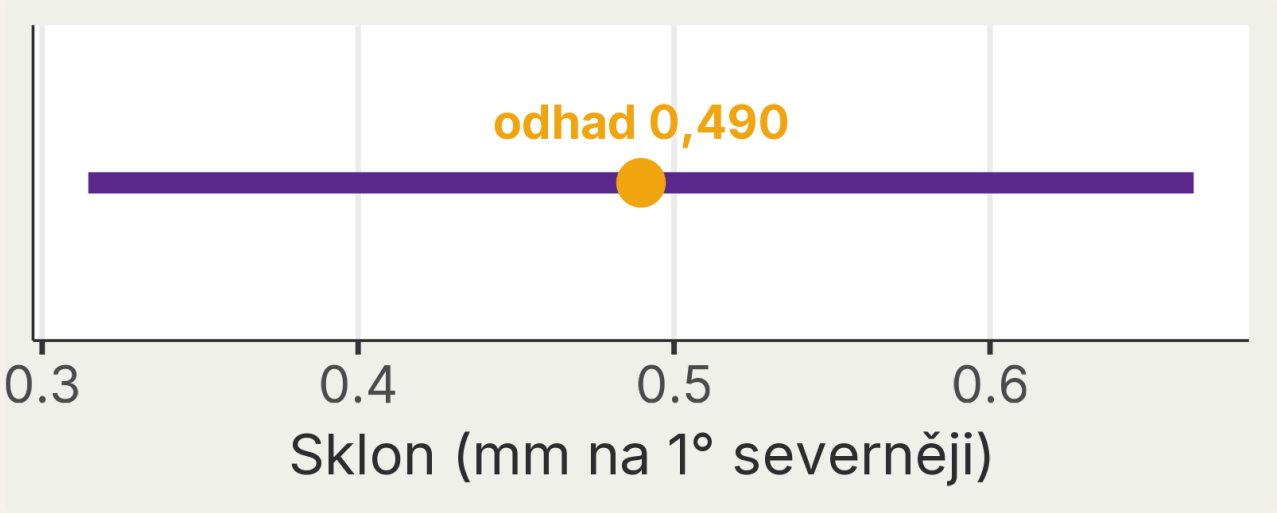


Výsledek se změnil, protože se změnila pozorování; předpokládaný biologický vztah zůstal stejný.

Pro názornost jsou z 200 simulací vybrány pohledy s nízkým, středním a vysokým odhadnutým sklonem.

Jak široké je naše okno do reality?

Odhad a 95% interval spolehlivosti



CO ODHADUJEME

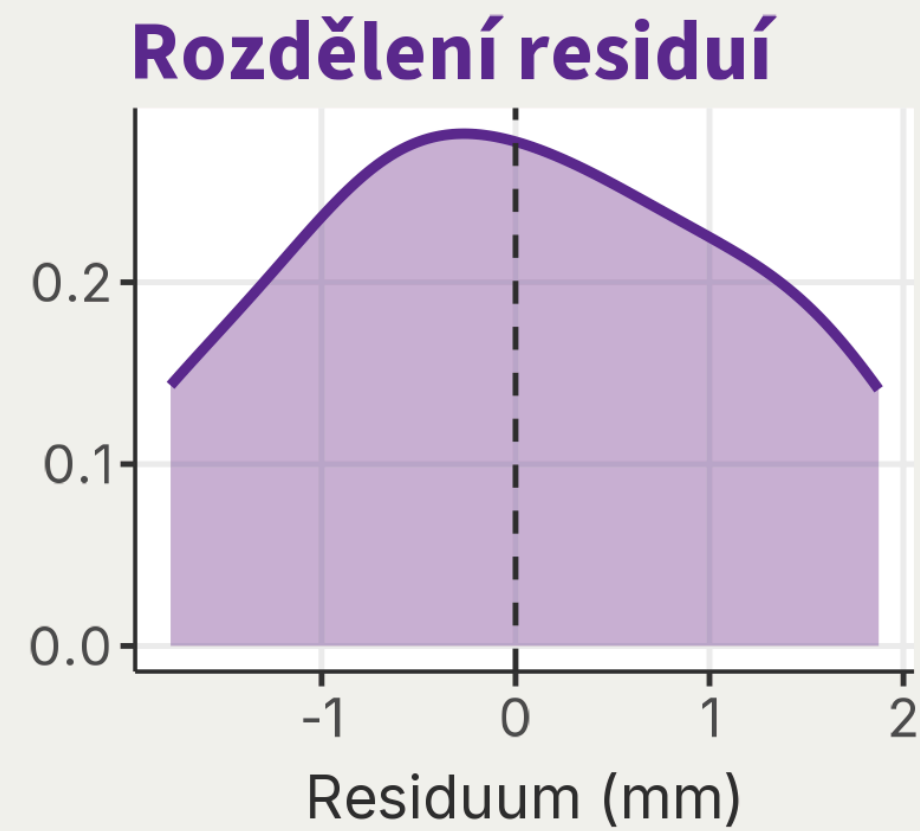
Sklon: **0,490 mm/°**

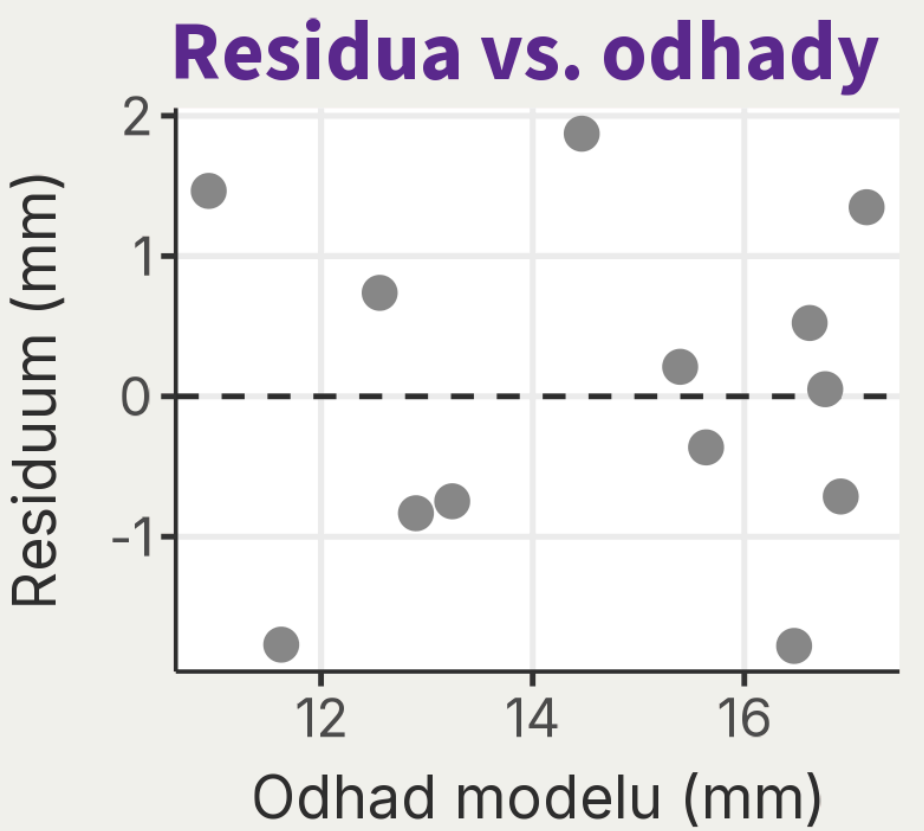
JAK PŘESNĚ

95% CI: **[0,315; 0,664] mm/°**

Pouze 13 lokalit a jednoduchý observační model.

Rychlá kontrola modelu





V residuích nevidíme nápadný problém, který by jednoduchý lineární model okamžitě zpochybnil.

Co už víme — a co stále nevidíme?

CO ZACHYTILA STUDIE

Kladný pozorovaný sklon

- jeho velikost v mm/°
- interval slučitelných efektů
- žádný nápadný problém v základní diagnostice

CO NEVIDÍME PŘÍMO

- skutečný populační sklon
- výsledek jiné možné studie
- příčinný biologický mechanismus
- zda by pozorovaný vztah nemohl vzniknout prostou náhodou

Co kdyby žádný trend nebyl?

Co bychom pozorovali bez trendu?

HYPOTETICKÁ REALITA
Skutečný systematický sklon je **přesně 0 mm/°**.
Lokality se přesto biologicky liší.

Musí být odhadnutý sklon v každé pozorované studii také přesně nulový?

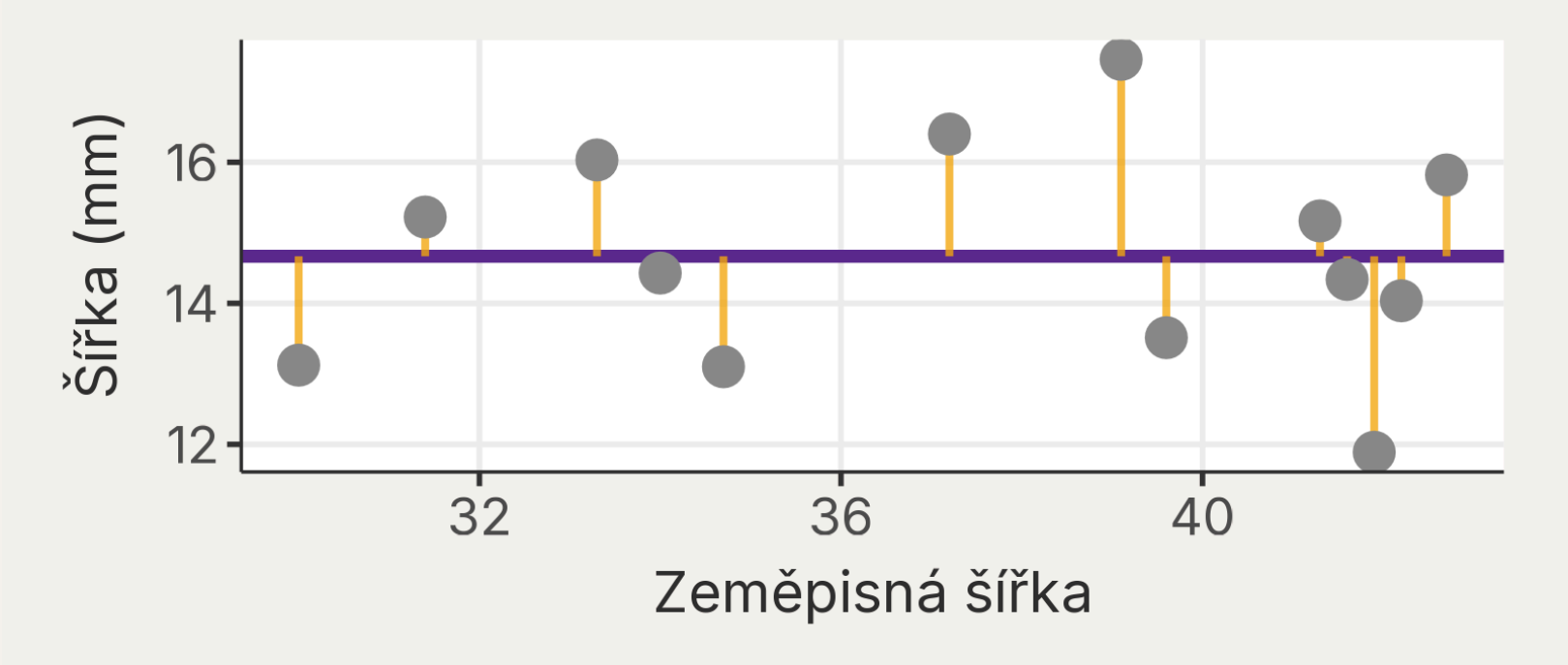
A) Ano, všechny přímky budou vodorovné.

B) Ne, sklony se budou náhodně lišit kolem nuly.

C) Ne, všechny budou kladné.

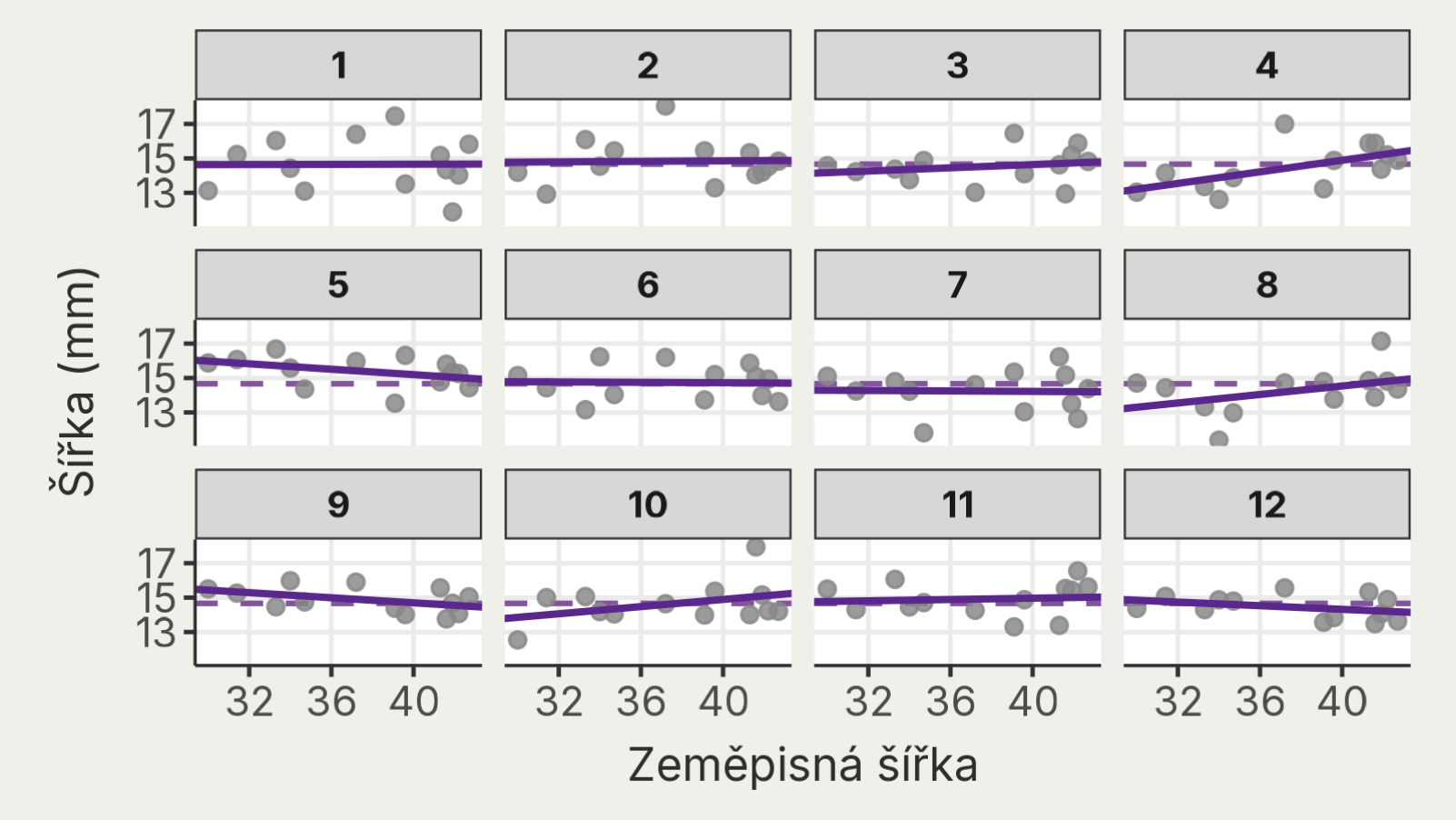
Jak náhoda vytvoří nenulový sklon?

pozorovaná velikost = **základní úroveň + 0 ×** **zeměpisná šířka** + **náhodné residuum**



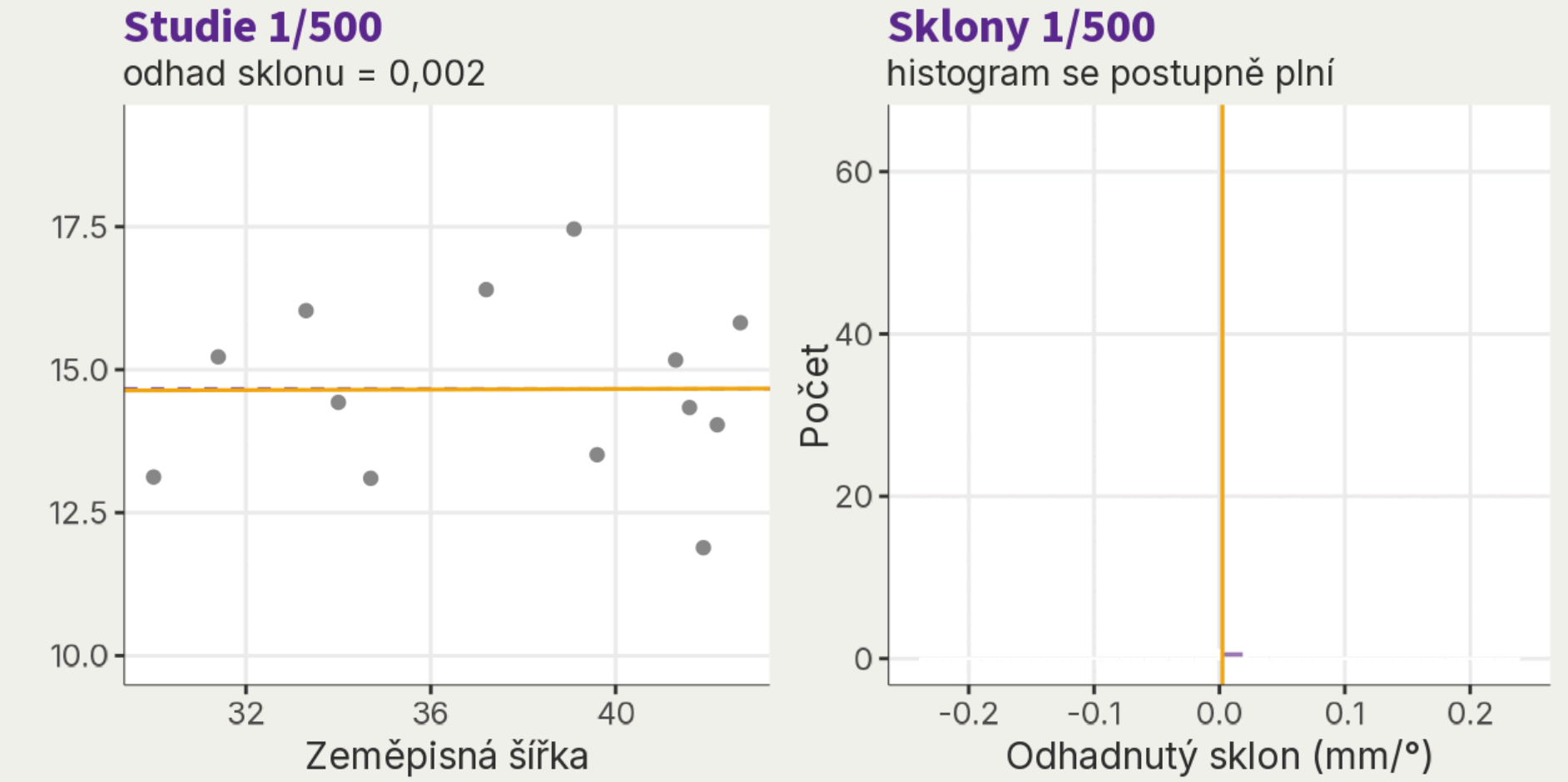
Systematický sklon zůstává nulový. Náhodná residua posouvají jednotlivé body nad a pod vodorovnou přímku.

Nulový trend, různé odhadnuté přímky



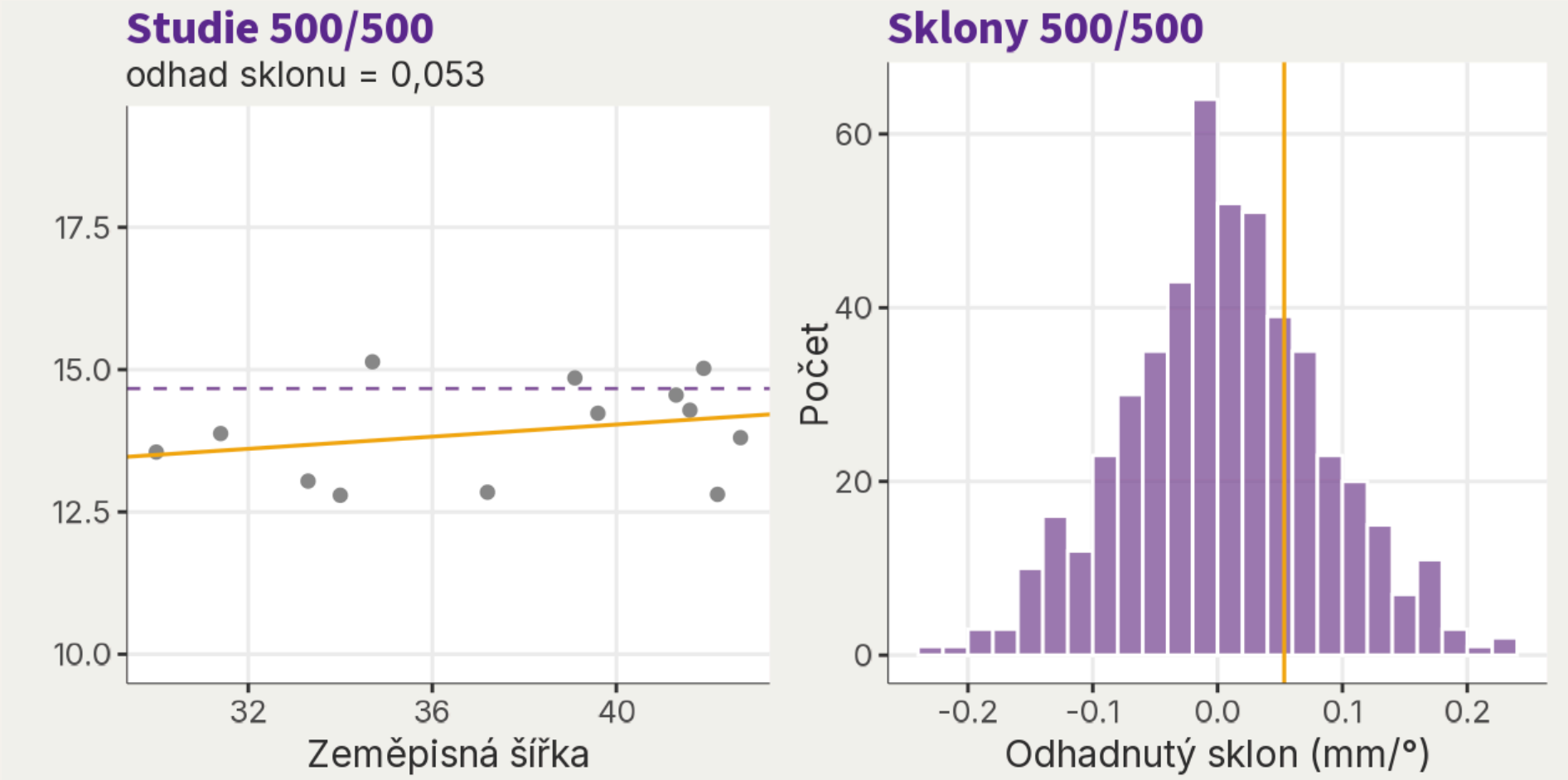
Nulová realita zůstala stejná. Měnila se pozorování, a proto se měnily i odhadnuté sklony.

Sledujte cestu jednoho náhodného sklonu



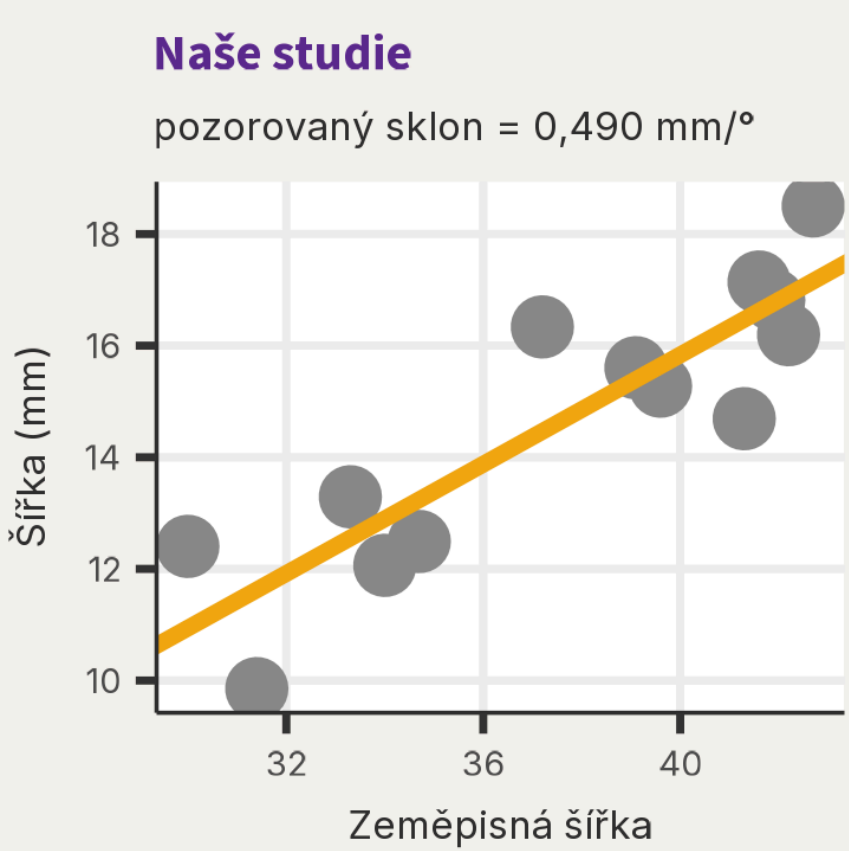
Nová náhodná residua → nový soubor bodů → nový sklon → další část histogramu.

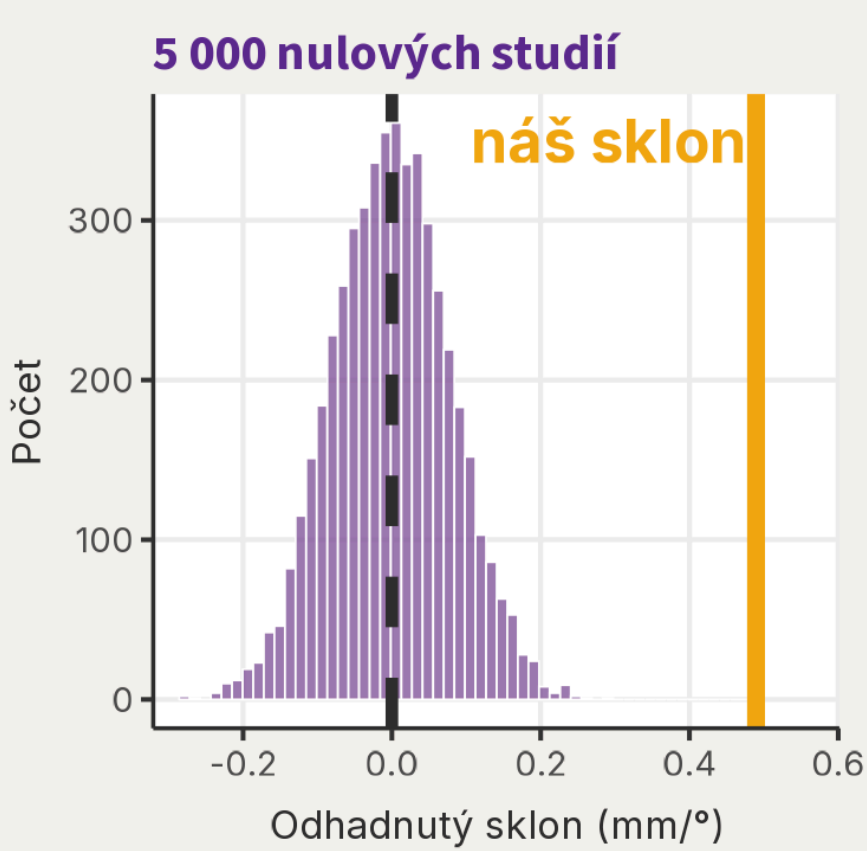
Po 500 nulových studiích



Kolem které hodnoty se sklony shlukují? Kam byste umístili náš pozorovaný sklon?

Stejný nulový svět, ale s naším pozorováním

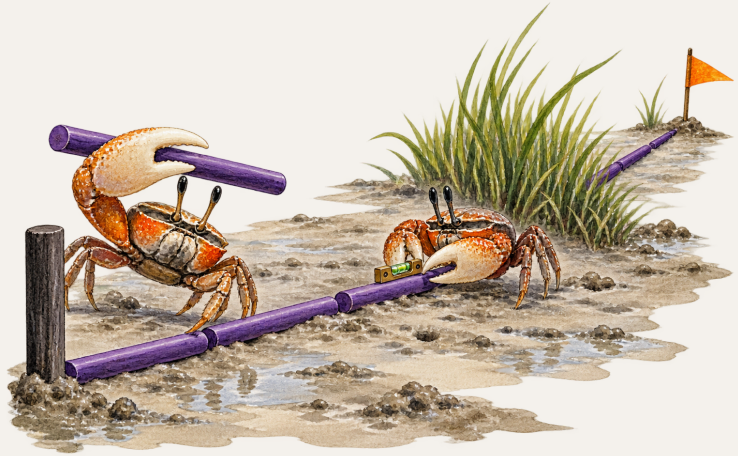




Nenulové sklony vznikají v nulovém světě běžně. Sklon tak velký jako náš je však mimořádně daleko od většiny simulací.

Jak vzdálené je naše pozorování od nuly?

	Hodnota
Pozorovaný efekt	0,490 mm/°
Efekt v nulovém světě	0 mm/°
Rozdíl	0,490 mm/°
Standardní chyba odhadu	0,0795 mm/°



Ilustrace vytvořená pomocí AI.

Kolikrát se standardní chyba vejde do rozdílu mezi pozorovaným a nulovým sklonem?

Vzdálenost ve standardních chybách

Vzdálenost v jednotkách standardní chyby (SE) = rozdíl mezi pozorovaným a nulovým sklonem ÷ standardní chyba (SE) sklonu

$$\frac{0{,}490 - 0}{0{,}0795} = 6{,}16$$

$$t = \frac{\text{pozorovaný sklon} - \text{nulový sklon}}{\text{standardní chyba (SE) sklonu}} \approx 6{,}16$$

Této standardizované vzdálenosti říkáme t-statistika.

Naše pozorování leží přibližně **6,16 standardních chyb od nuly**.

Kolik plochy zůstává za $|t|$?

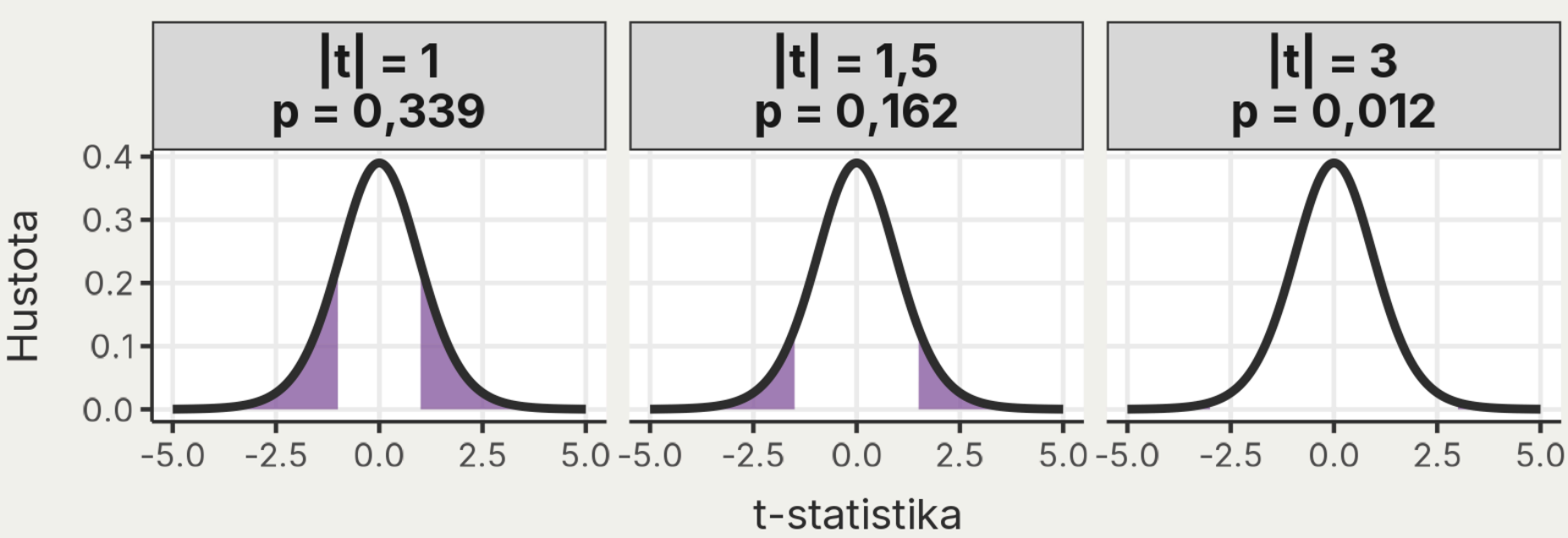


Křivka je **t-rozdělení**: hodnoty t, které by vznikaly v nulovém světě. Celá plocha pod ní je 1.

P-HODNOTA

Pravděpodobnost získat v nulovém světě t alespoň tak vzdálené od nuly jako naše. U krabů $p = 0{,}000071$.

Co udělá větší hodnota $|t|$?

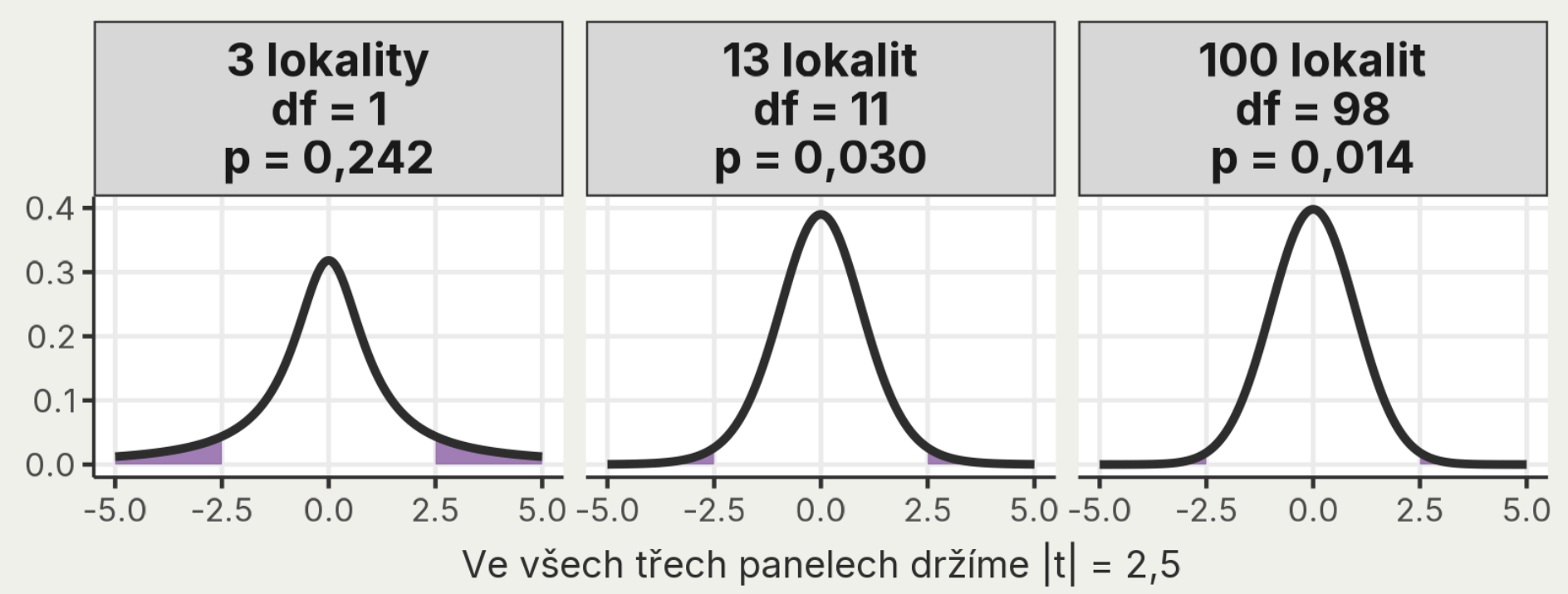


Ve všech třech panelech stejný počet lokalit jako u krabů (13)

Když se $|t|$ vzdaluje od nuly, co se stane s fialovou plochou?

Čím dál leží t od nuly, tím menší část nulového rozdělení zůstává stejně daleko nebo ještě dál. Tím menší je p-hodnota.

Více lokalit při stejném $|t|$



Při stejném $|t| = 2{,}5$: ve které studii zůstane ve chvostech nejméně plochy?

Přímka odhaduje dva koeficienty, intercept a sklon, a proto platí $df = n - 2$, u krabů $13 - 2 = 11$. Více nezávislých pozorování znamená tenčí chvosty a menší p-hodnotu.

Která věta čte p-hodnotu správně?

A) Pravděpodobnost, že skutečný sklon je nulový, je 0,000071.

B) V nulovém světě by výsledek alespoň tak extrémní jako náš byl velmi neobvyklý.

C) Biologický efekt má velikost 0,000071.

Přesvědčte souseda, proč jste vybrali svou možnost.

Co nám malá p-hodnota řekne — a co ne?

CO ŘÍKÁ

Naše pozorování je velmi obtížně slučitelné s přesně nulovým systematickým sklonem a zvoleným modelem.

CO NEŘÍKÁ

- že je biologická hypotéza pravdivá
- že je efekt velký nebo důležitý
- že model nemá omezení
- jaký mechanismus vztah vytváří

P-hodnota hodnotí pozorování v jednom hypotetickém světě. Není to stroj na pravdu.

Nulová hypotéza pojmenuje nulový svět



Ilustrace vytvořená pomocí AI.

H_0: NULOVÁ HYPOTÉZA

Pro kraby říká:

$$H_0 : \text{ systematický sklon} = 0 \text{ mm}/^\circ$$

H_A: ALTERNATIVNÍ HYPOTÉZA

$$H_A : \text{ systematický sklon} \neq 0 \text{ mm}/^\circ$$

Nulová hypotéza netvrdí, že se lokality neliší. Tvrdí jen, že se průměrná velikost se zeměpisnou šířkou systematicky lineárně nemění.

Kdy musíme udělat rozhodnutí?

Někdy nestačí popisovat — musíme jednat

Monitoring zachytil jednoho jedince možného invazního druhu.

Zasáhnout, nebo pokračovat v pozorování? Jak velké riziko falešného poplachu jsme ochotni přijmout?



Ilustrace vytvořená pomocí AI.

Kolik falešných poplachů jsme ochotni připustit?

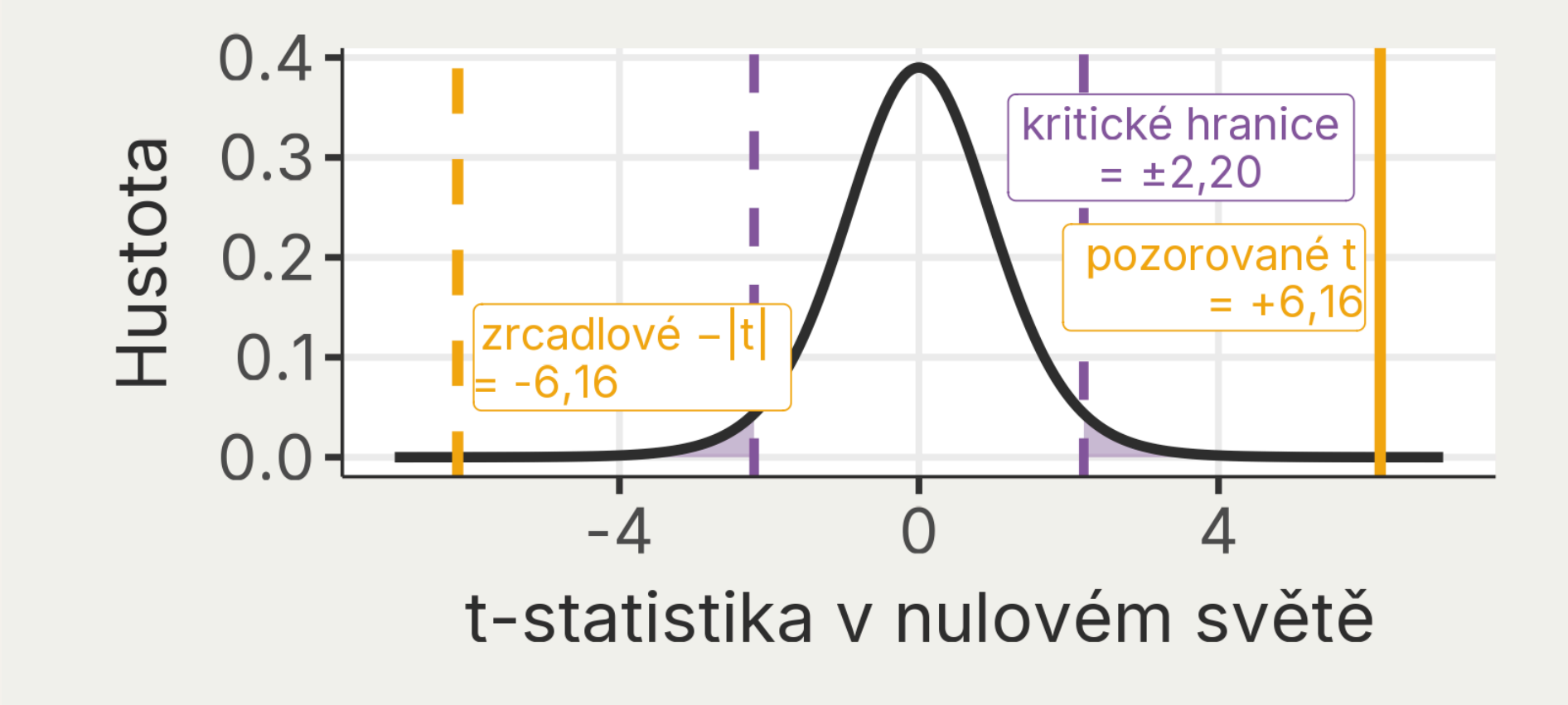
Kdyby nulová hypotéza platila, jak často jsme ochotni ji **omylem** zamítnout?

α: HLADINA VÝZNAMNOSTI

Předem zvolená přijatelná míra chyby. $\alpha = 0{,}05$ znamená 5 %.

Pět procent je běžná konvence. Podle důsledků chyby můžeme zvolit i přísnější pravidlo.

Kde začíná krajních 5 % nulového světa?



Pro $df = 11$ fialové hranice oddělují prostředních 95 % od chvostů o celkové ploše $\alpha = 0{,}05$. Naše $|t| = 6{,}16$ leží daleko za nimi.

Rozhodovací pravidlo volíme předem

ROZHODOVACÍ PRAVIDLO

Porovnáme p-hodnotu s předem zvolenou hladinou α.

KDYŽ $p > \alpha$

Nulovou hypotézu zamítnout **nemůžeme**.

KDYŽ $p \leq \alpha$

Nulovou hypotézu zamítneme. Výsledek nazveme **statisticky významný** neboli **statisticky signifikantní**.

Před počítači: $|t|$ vs. kritická hodnota z tabulky pro dané df. V R: `t value` a `Pr(>|t|)`.

Co když se spleteme?

Chyba I. typu



Falešně pozitivní: zamítneme H_0, přestože platí.

Chyba II. typu



Falešně negativní: nezamítneme H_0, přestože neplatí.

Invazní druh: která chyba je která?

$$H_0 : \text{invazní druh v oblasti není přítomen}$$

Realita	Průzkum: „nepřítomen“	Průzkum: „přítomen“
Druh není přítomen	správné rozhodnutí	**A: zbytečný zásah**
Druh je přítomen	**B: přehlédnutá invaze**	správné rozhodnutí

Zařaďte A a B jako chybu I. nebo II. typu. Která může být závažnější?

A = chyba I. typu.
B = chyba II. typu.
Závažnost závisí na biologických a praktických důsledcích.

Co tedy můžeme tvrdit?

„p < 0,05, proto je hypotéza pravdivá“

Co je na této větě špatně?

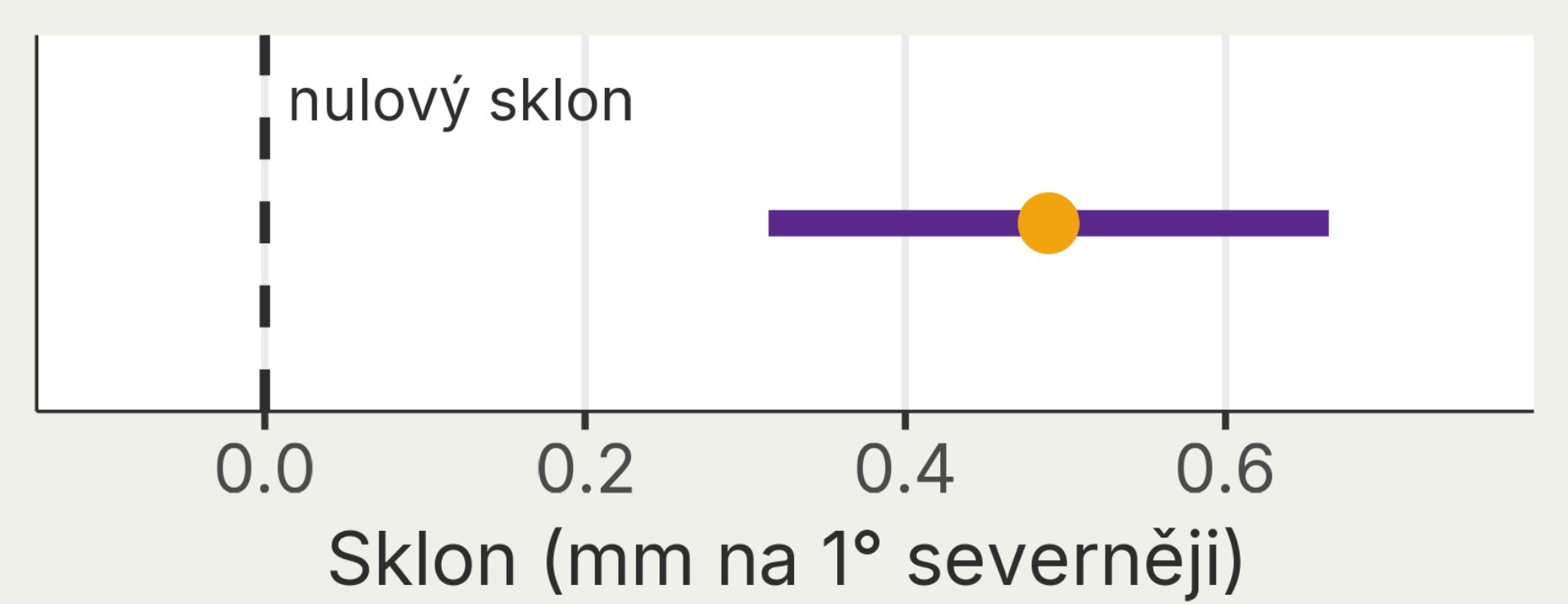
CO TVRDÍ NEPRÁVEM

P-hodnota neudává pravděpodobnost pravdy a test hypotézu nedokazuje.

CO VE VĚTĚ CHYBÍ

Efekt, jednotky, biologický význam, nejistota a omezení.

Ke stejné nule se můžeme vrátit intervalem



95% interval **[0,315; 0,664] mm/°** neobsahuje nulu. Při $\alpha = 0{,}05$ proto vede ke stejnému rozhodnutí jako p-hodnota a navíc ukazuje slučitelné velikosti efektu.

Napište výsledek pro biologa

Napište 4–6 vět. Začněte tím, **co se mění a o kolik**. Poslední větou uveďte jednu důležitou limitaci studie.

Přesvědčte souseda, že vaše pořadí důkazů dává smysl.

KARTA DŮKAZŮ

- 13 slanisek
- pozorovaný sklon 0,490 mm/°
- 95% CI [0,315; 0,664] mm/°
- $t = 6{,}16$, $df = 11$, $p = 0{,}000071$
- observační studie a jednoduchý lineární model
- možné místní a prostorové rozdíly mezi lokalitami

Jedna možná odpověď

Efekt: Na 13 sledovaných slaniskách se průměrná šířka krunýře směrem k severu zvětšovala asi o **0,490 mm na 1°**.
Nejistota: 95% interval spolehlivosti byl **[0,315; 0,664] mm/°**.
Biologický význam: Přes celé sledované pobřeží to odpovídá změně asi 6,2 mm, tedy znatelnému rozdílu ve velikosti krabů.
Statistická významnost: Výsledek je při $\alpha = 0{,}05$ statisticky významný ($t = 6{,}16$, $df = 11$, $p = 0{,}000071$).
Limitace: Studie podporuje kladný vztah na sledovaných lokalitách, ale sama neodhaluje příčinu ani přesný vztah v celé populaci.

Co nám studie ukázala?

HLAVNÍ ZÁVĚR

Na 13 sledovaných slaniskách byli krabi směrem na sever v průměru větší. Odhadnutý vztah činil **0,490 mm na 1°** a kladný trend směrem k severu je statisticky významný.

CO ZŮSTÁVÁ OTEVŘENÉ

- přesná velikost vztahu v dalších lokalitách
- příčinný biologický mechanismus
- vliv místních a prostorových podmínek



Ilustrace vytvořená pomocí AI.

Nejistota neznamená, že studie nic neukázala. Vymezuje, jak daleko může náš závěr bezpečně sahat.

A co sýrový Měsíc?



Prokázal by statistický test, že Měsíc není ze sýra?

Ne. Ukázal by, jak slučitelná jsou pozorování s jedním přesně popsaným sýrovým světem.

Závěr

Platí stejná logika i pro dvě skupiny?

TATO LEKCE

Číselný prediktor: zeměpisná šířka.
Efekt: **sklon**.

NÁSLEDUJÍCÍ LEKCE

Kategoriální prediktor: například kontrola a zásah.
Efekt: **rozdíl mezi skupinami**.

Co z této logiky zůstane stejné, když místo trendu budeme porovnávat dvě biologické skupiny?

Shrnutí

- Studie zachycuje omezené pozorování reality.
- Výsledek začíná velikostí efektu a jeho nejistotou.
- Nulová hypotéza popisuje jeden přesně zadaný svět; t-statistika měří vzdálenost od něj ve standardních chybách.
- P-hodnota popisuje, jak neobvyklá jsou data v nulovém světě, ne pravděpodobnost hypotézy.
- Rozhodnutí může být chybné: chyba I. a II. typu.

Efekt → biologický význam → nejistota → p-hodnota → limitace závěru

Co potřebujete vidět, než napíšete „p < 0,05, takže severní krabi jsou větší“?

Jaké tři otázky položíte autorovi, který vám ukáže jen tento titulek?

ŘEKNĚTE SOUSEDOVI JEDNOU VĚTOU

Kterou z otázek byste položili jako první, a proč?