BIOSTATISTIKA

L10

Jak rychle roste počet druhů rostlin s plochou ostrova?



Tvar rozdělení, transformace a nelinearita

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

Ilustrace vytvořená pomocí AI. Družicový snímek: NASA/GSFC (volné dílo).

Kolikrát víc druhů rostlin má desetkrát větší ostrov?



Na větších ostrovech roste více druhů rostlin. Ale kolikrát více?

- **A** · asi 1,5krát
- **B** · asi 2,5krát
- **C** · asi 5krát
- **D** · 10krát

Snímek: J. Descloitres, MODIS Rapid Response Project, NASA/GSFC (volné dílo), 12. 3. 2002

Co si pamatujete z minulé lekce?

Odpovězte samostatně na tři krátké otázky. Společné odpovědi projdeme po hlasování.

Proč model Různé sklony nemá nejvyšší adjustované R^2, přestože má nejvyšší R^2?

Tři kandidátní modely pro výšku stromů

odezva: vyska_stromu_m · stejných 63 stromů v každém modelu

Model	k	R^2	adjustované R^2
Společná přímka	2	0,768	0,764
Posun mezi druhy	4	0,793	0,782
Různé sklony	6	0,797	0,779

- Adjustované R^2 je podíl správně předpovězených stromů.
- Model Různé sklony byl odhadnut z jiných stromů.
- Dva sklony navíc zlepšily shodu s daty příliš málo.
- Adjustované R^2 je u složitějšího modelu vždy nižší.

SPRÁVNÁ ODPOVĚĎ

: Dva sklony navíc zlepšily shodu s daty příliš málo.

R^2 po přidání koeficientů neklesne; adjustované R^2 za každý koeficient navíc penalizuje.

Model Posun mezi druhy má ΔAIC = 0. Co to znamená?

AIC tří kandidátních modelů

odezva: vyska_stromu_m · stejných 63 stromů v každém modelu

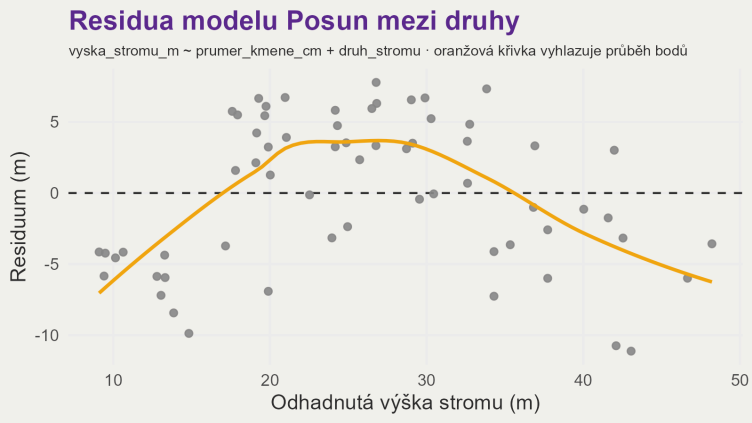
Model	AIC	ΔAIC
Společná přímka	398,64	3,20
Posun mezi druhy	395,44	0,00
Různé sklony	398,27	2,83

- Má nejnižší AIC z těchto tří kandidátů odhadnutých ze stejných 63 stromů.
- Jeho residua jsou nulová, takže model předpovídá výšky bez chyby.
- Je to pravdivý model růstu stromů.
- Jeho AIC můžeme porovnat i s AIC modelu odhadnutého z jiných dat.

SPRÁVNÁ ODPOVĚĎ

: Má nejnižší AIC z těchto tří kandidátů odhadnutých ze stejných 63 stromů.
ΔAIC = 0 označuje nejnižší AIC v této sadě kandidátů. Neříká nic o pravdivosti modelu; AIC porovnáváme jen při stejné odezvě a stejných pozorováních.

Co prozrazuje oblouk v residuích modelu Posun mezi druhy?



- Model je v pořádku, protože residua jsou v průměru nulová.
- Stromy jednoho druhu jsou si podobnější než stromy různých druhů.
- Některé stromy jsou chybně změřené a je potřeba je vyřadit.
- Přímky nezachycují tvar vztahu: krajní stromy model přeceňuje, střední podceňuje.

SPRÁVNÁ ODPOVĚĎ

: Přímky nezachycují tvar vztahu: krajní stromy model přeceňuje, střední podceňuje.
Residua pod nulou = model výšku přeceňuje. Oblouk ukazuje zakřivený vztah, který přímky nezachytí.

Výsledky učení

Po lekci dokážete:

- popsat tvar rozdělení proměnné nebo residuí z histogramu a Q–Q grafu a posoudit, zda je pro lineární model podstatný;
- rozlišit transformace, které mění jen měřítko (jednotky, centrování, z-skóre), od těch, které mění tvar (logaritmus, odmocnina);
- rozpoznat zakřivený vztah v datech a navrhnout jednoduché řešení;
- porovnat několik biologicky smysluplných tvarů vztahu jako kandidátní modely;
- vysvětlit, jak transformace mění interpretaci koeficientů.

Tvar rozdělení

Jeden řádek tabulky představuje jeden ostrov



U každého ostrova Galapág známe

- **počet druhů** cévnatých rostlin;
- **plochu** ostrova (km²);
- **nejvyšší bod** (m n. m.);
- počet endemitů a vzdálenosti k sousedním ostrovům.

Botanici druhy spočítali na ostrovech od skalek po obří sopky.

Ilustrace vytvořená pomocí AI.

Data: M. P. Johnson & P. H. Raven (1973), *Science* 179: 893–895 · balíček `faraway` 1.0.9 (`galamiss`) · GPL

Co v tabulce chybí a co je neobvyklé?

```
data_ostrovy <- read.csv(
  file = here::here("data", "druhy_galapagy.csv")
)

colSums(x = is.na(data_ostrovy))
```

```
                    ostrov                pocet_druhu
                         0                          0
                plocha_km2            nejvyssi_bod_m
                         0                          6
 vzdalenost_santa_cruz_km    plocha_sousedniho_km2
                         0                          0
```

```
range(x = data_ostrovy$plocha_km2)
```

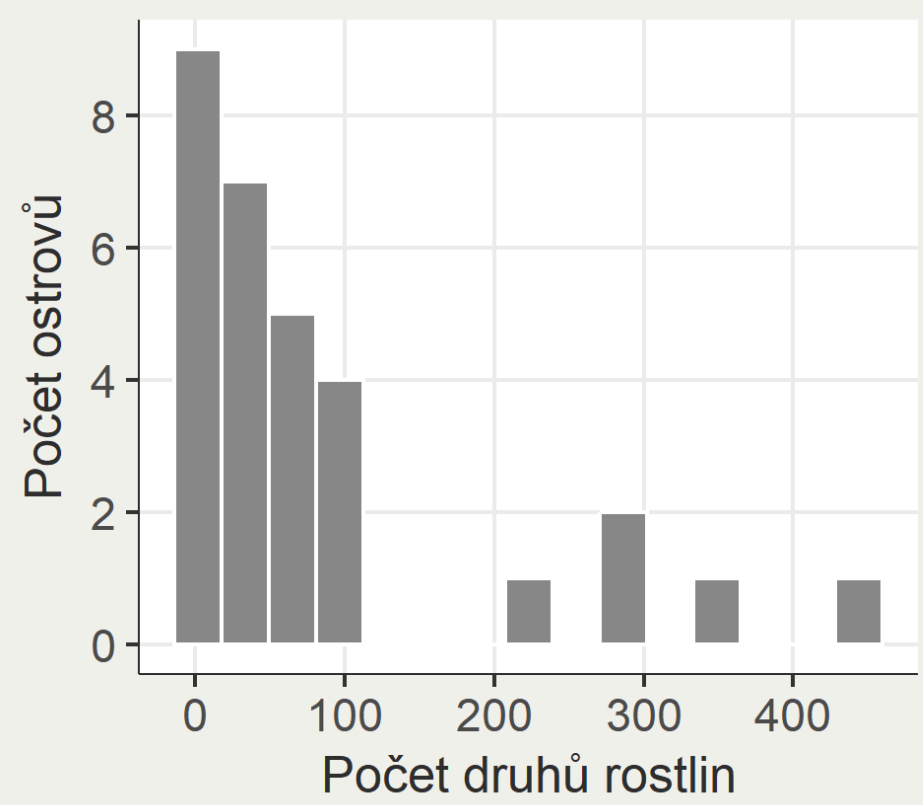
```
[1]    0.01 4669.32
```

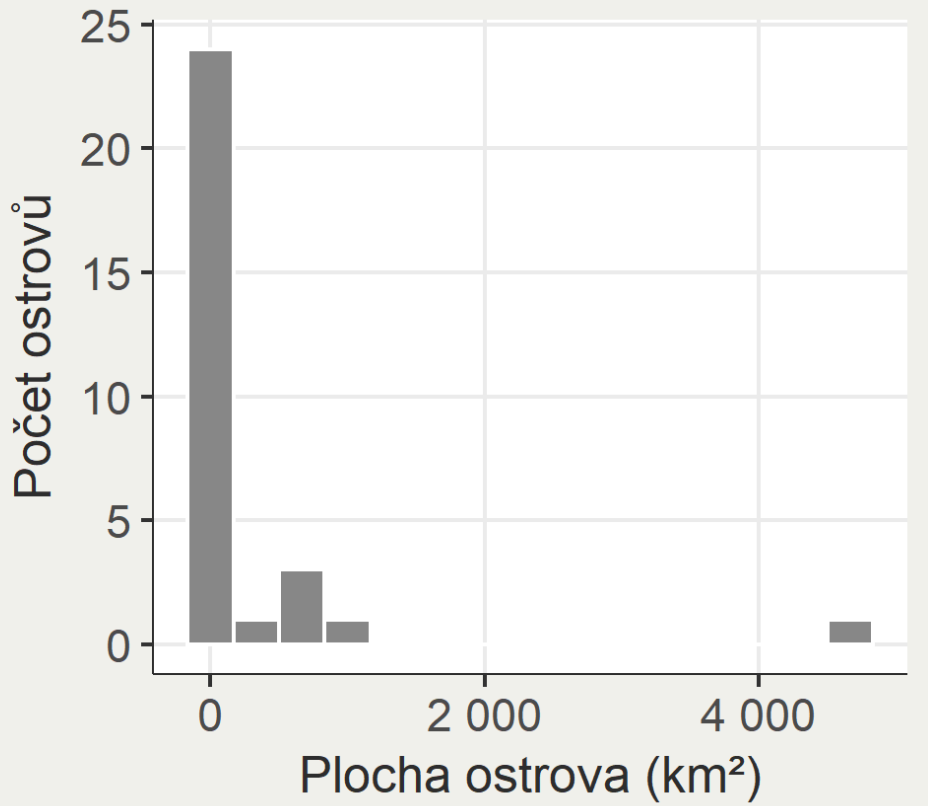
Které hodnoty by nám mohly dělat potíže?

30 ostrovů

- nejvyšší bod chybí u 6 ostrovů;
- plocha od 0,01 do 4 669 km² (Isabela);
- skutečné nuly: vzdálenost ostrova Santa Cruz k sobě samému, Daphne Minor bez endemitů.

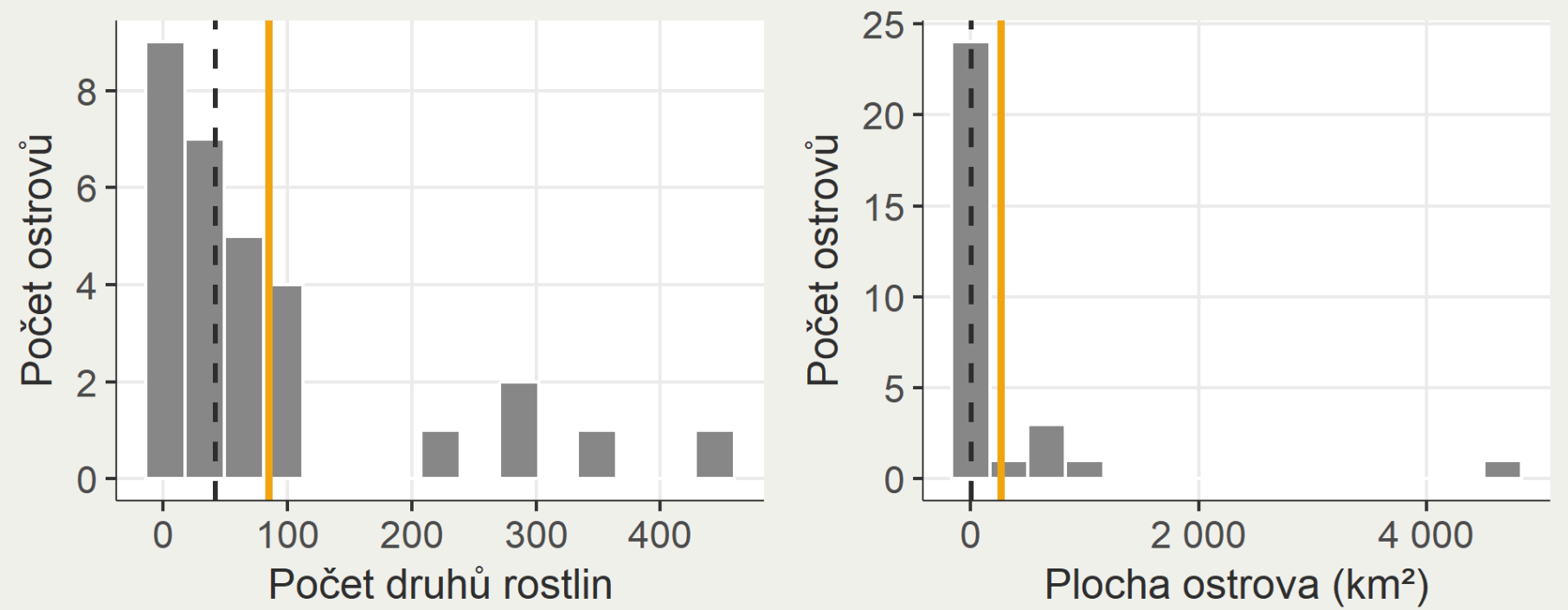
Jak vypadají počty druhů a plochy ostrovů?





Ve dvojici: Kde leží většina ostrovů? Co s každým obrázkem dělá několik největších ostrovů?

Dlouhý ocas vpravo táhne průměr od mediánu

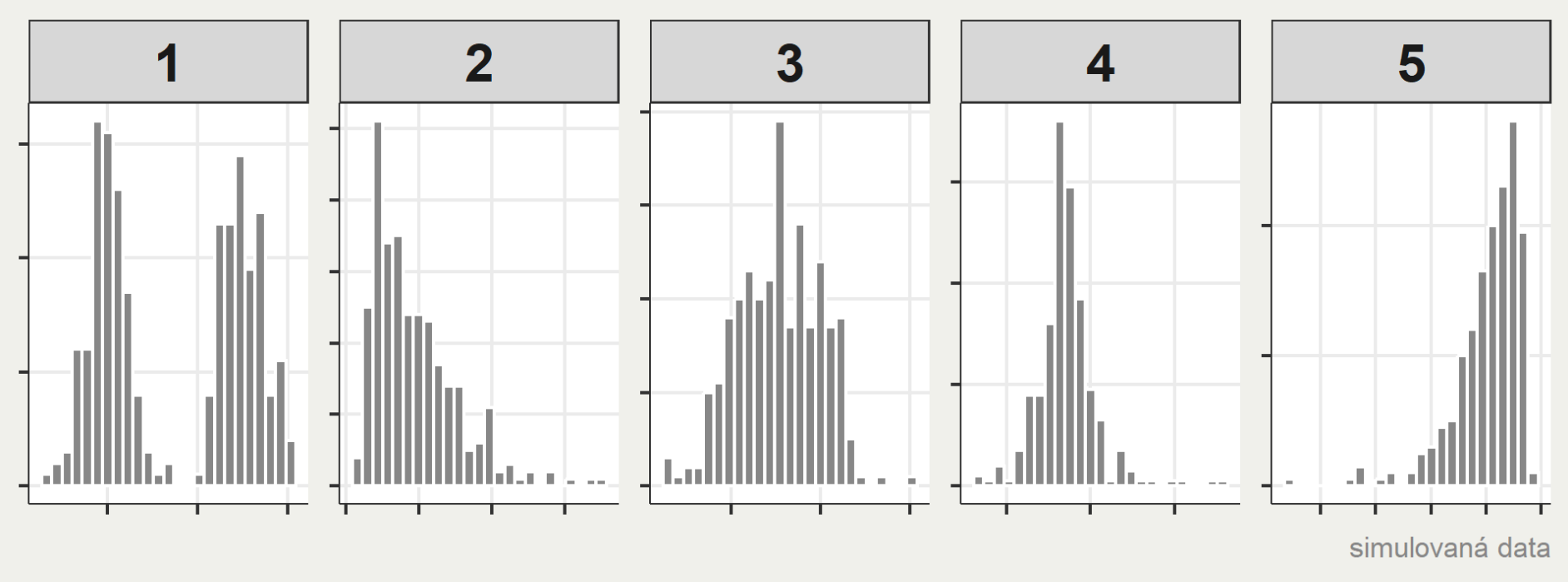


Druhy: průměr 85,2 · medián 42

Plocha (km²): průměr 261,7 · medián 2,59

> Rozdělení s dlouhým ocasem vpravo je **pravostranně šikmé** (*right-skewed*).

Které slovo patří ke kterému tvaru?



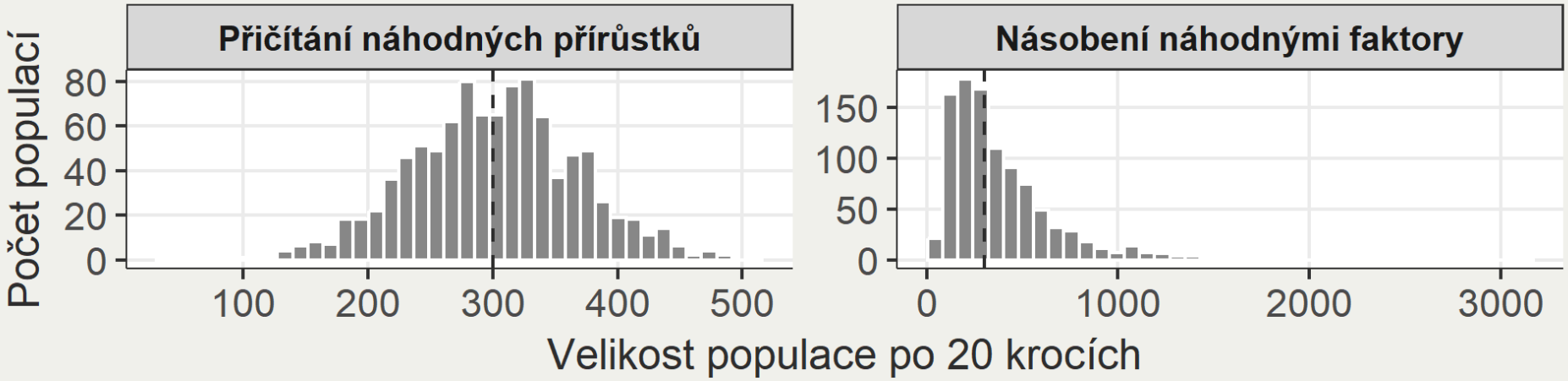
souměrné · pravostranně šikmé · levostranně šikmé · těžké konce · dva vrcholy

Ve dvojici: Přiřaďte každému panelu jedno slovo.

1 dva vrcholy · 2 pravostranně šikmé · 3 souměrné · 4 těžké konce · 5 levostranně šikmé

Proč je tolik biologických veličin pravostranně šikmých?

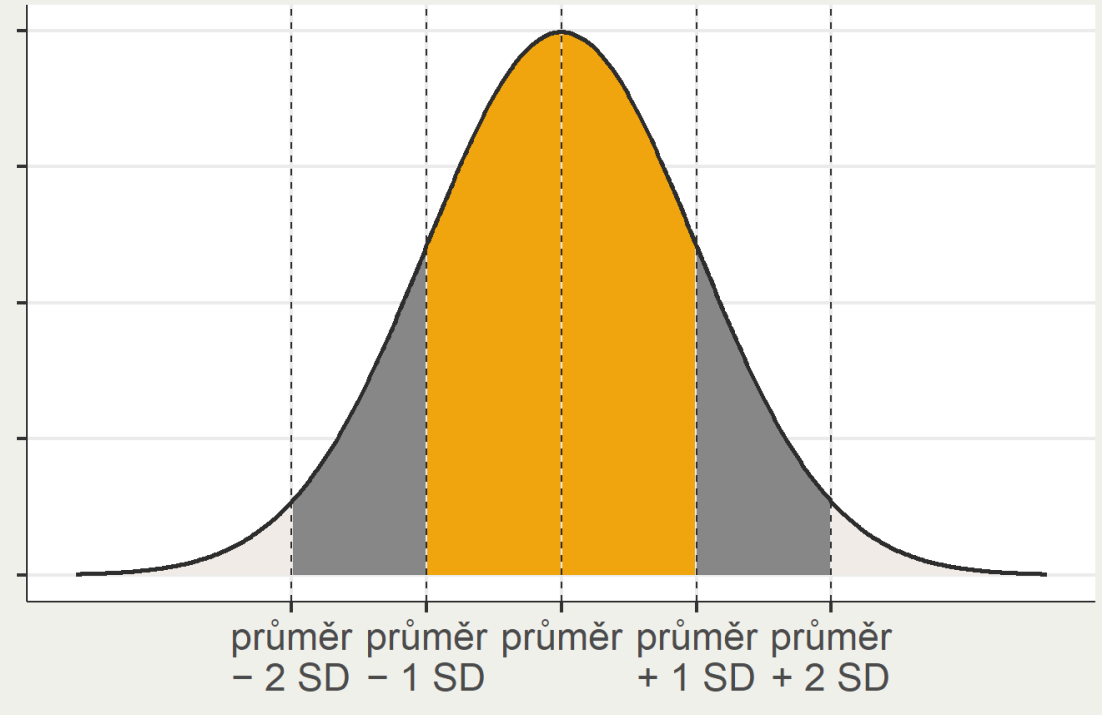
Plocha ostrova ani počet druhů nemohou klesnout pod nulu. Jak velké ale mohou být? A co se stane, když populace roste o **procenta**, ne o stejný počet jedinců?



simulovaná data

Hranice v nule a **růst násobením** vytvářejí dlouhý ocas vpravo — vlastnost biologie, ne chyba dat.

Kolik hodnot leží do jedné směrodatné odchylky od průměru?



Jaká část hodnot leží v oranžovém pásmu?

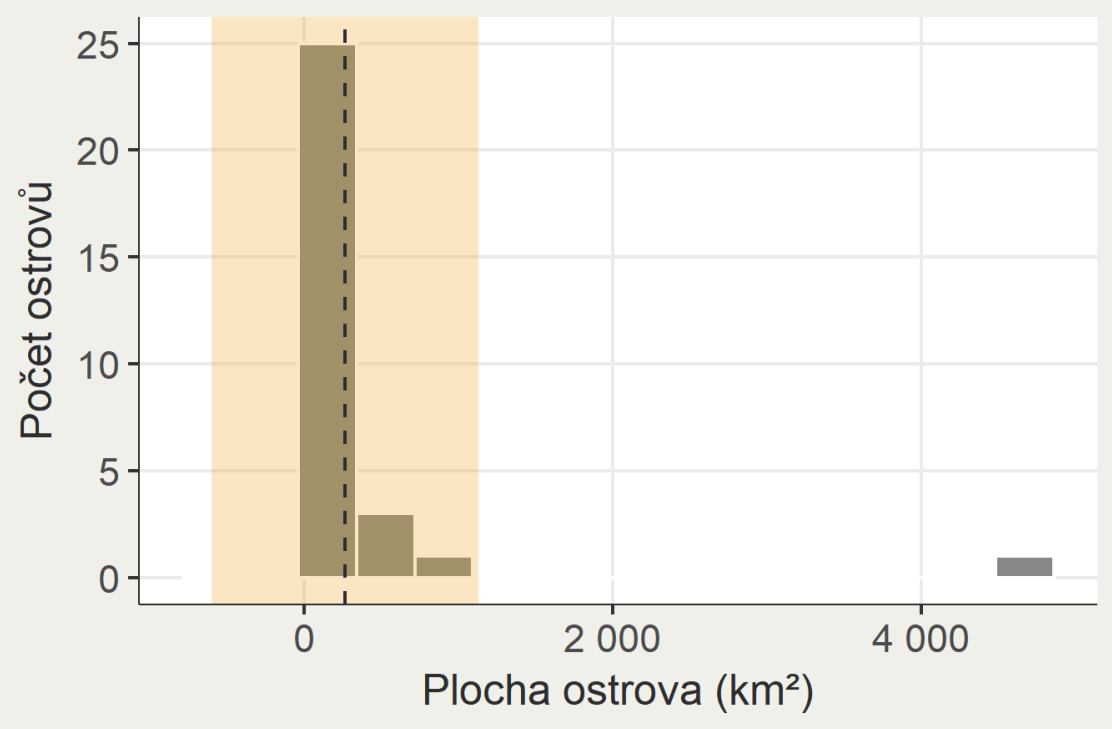
A · asi polovina

B · asi dvě třetiny

C · asi 95 %

Takový zvon je **normální rozdělení** (*normal distribution*): do 1 SD leží asi 68 %, do 2 SD asi 95 % a do 3 SD asi 99,7 % hodnot.

Platí pravidlo 68 % i pro plochy ostrovů?



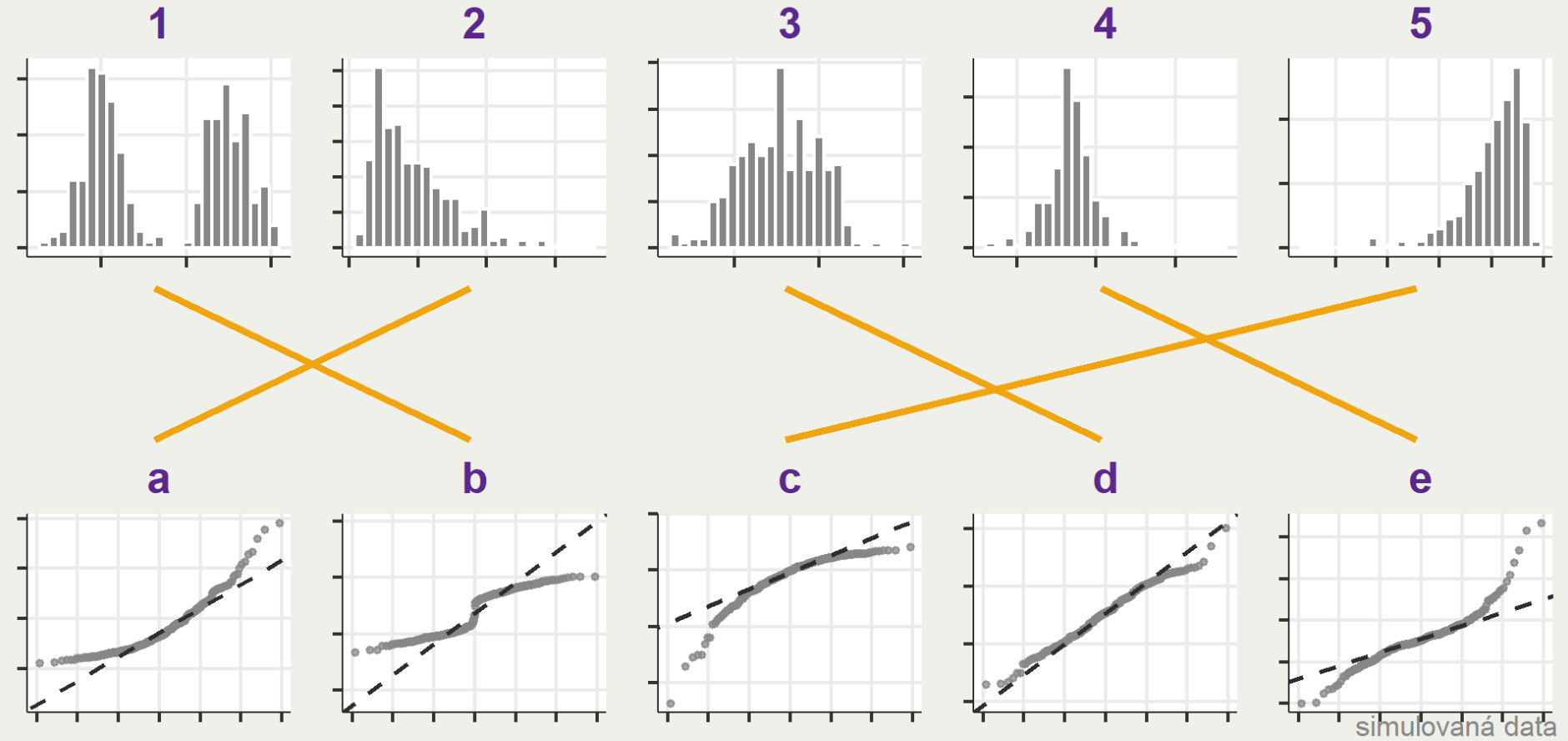
Ve dvojici: Pás ukazuje průměr ± 1 SD plochy. Co je na něm divné?

Průměr − SD = -602 km²: záporná plocha.

V pásu leží 97 % ostrovů, ne asi 68 %.

Normální rozdělení je vzor pro srovnání, ne vlastnost, kterou musí mít biologická data.

Jak se dlouhý ocas pozná v Q–Q grafu?



Ve dvojici: Který Q–Q graf a–e patří ke kterému histogramu 1–5?

Nápověda: dlouhý ocas vpravo zvedá body vpravo nad přímku.

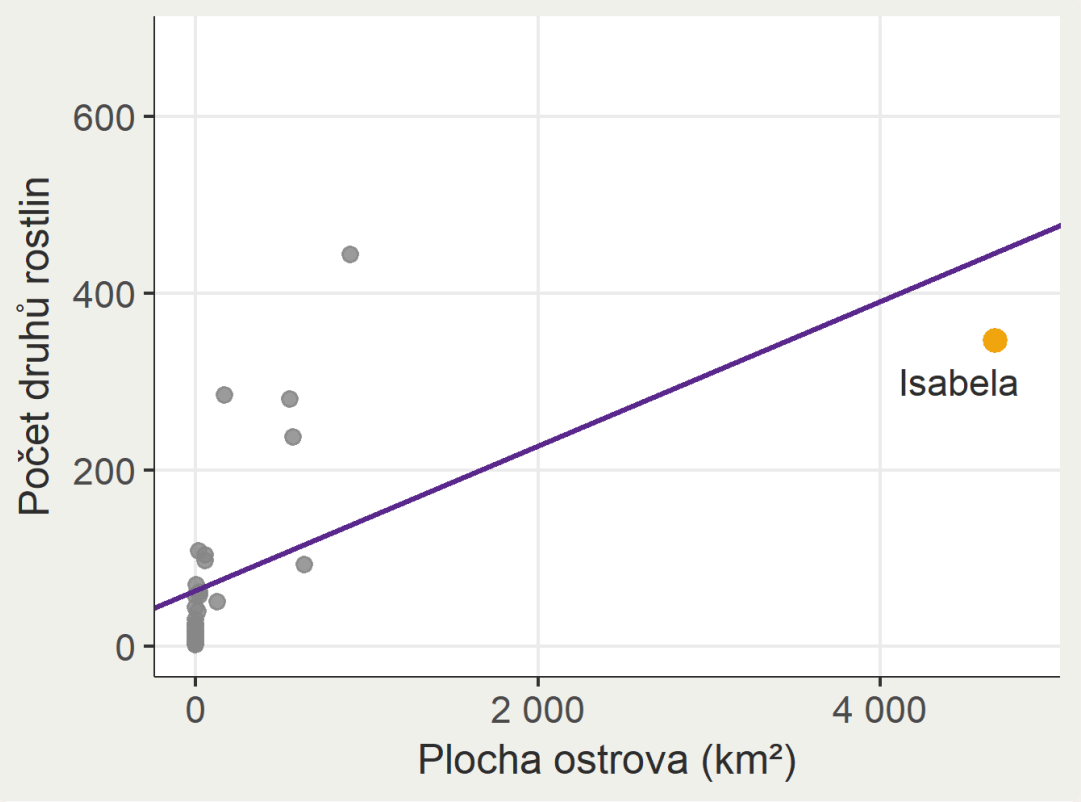
Sedí na ostrovy přímka?

Jak počet druhů popíše přímka?

```r
mod_primka <- lm(
  formula = pocet_druhu ~ plocha_km2,
  data = data_ostrovy
)

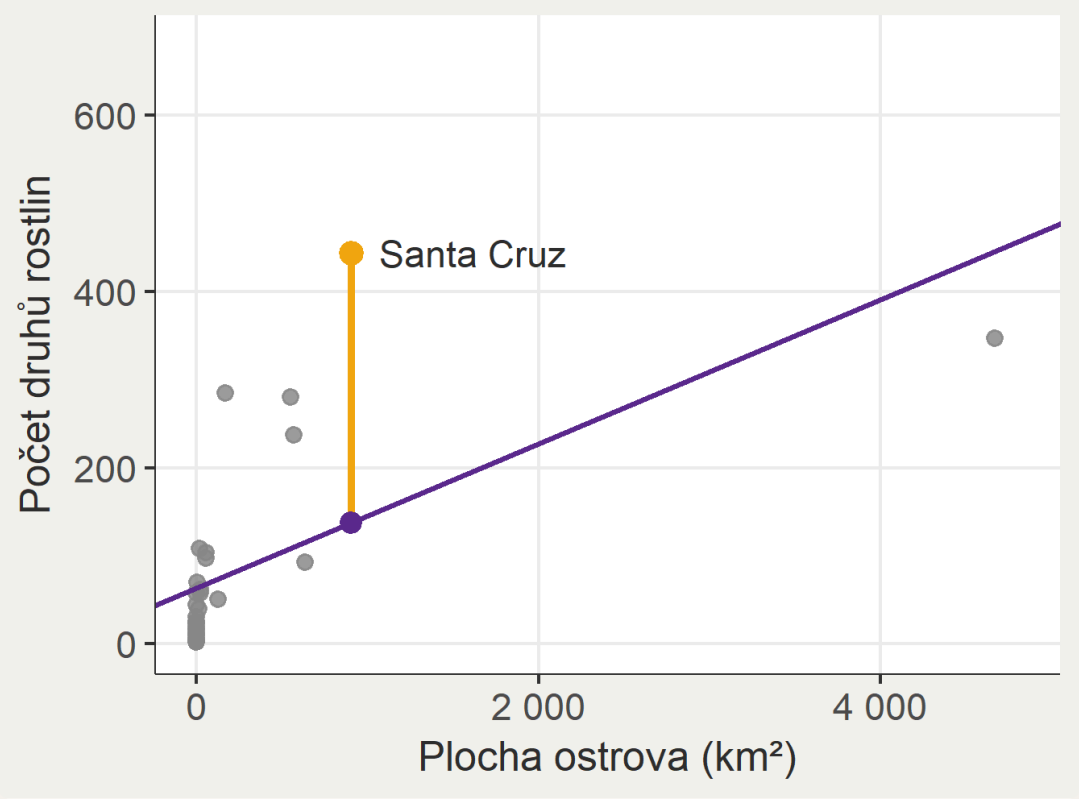
coef(object = mod_primka)
```

```
(Intercept)  plocha_km2
63.78286147  0.08196317
```



Kde leží většina ostrovů? Který ostrov o sklonu přímky rozhoduje nejvíc?

Z čeho se skládá počet druhů jednoho ostrova?



Santa Cruz: 444 druhů

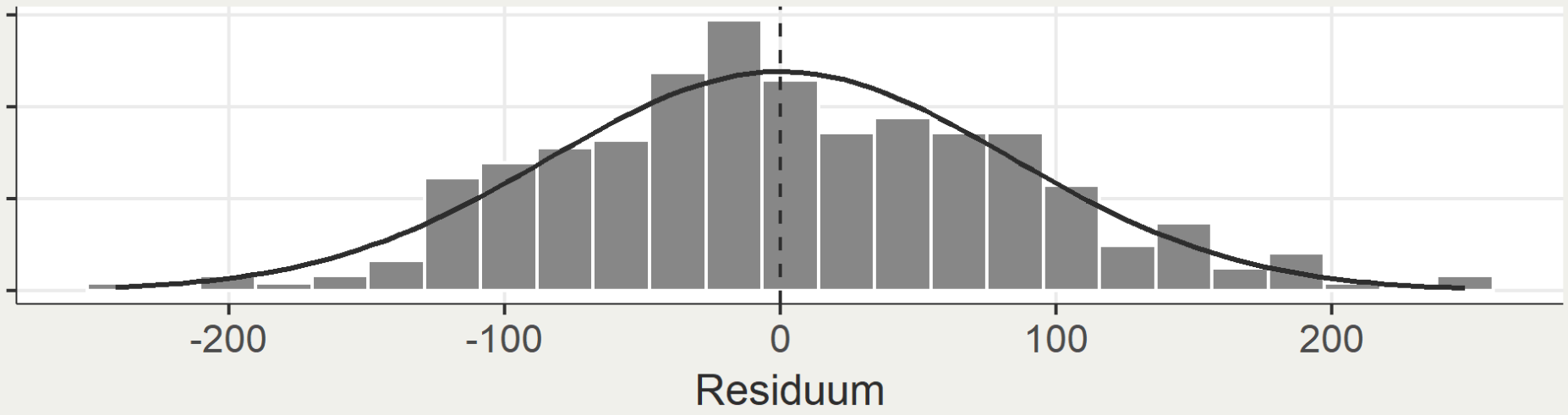
= odhad přímky 138

+ residuum 306

> Každý ostrov = přímka + residuum.

Jak by vypadala residua, kdyby přímka seděla?

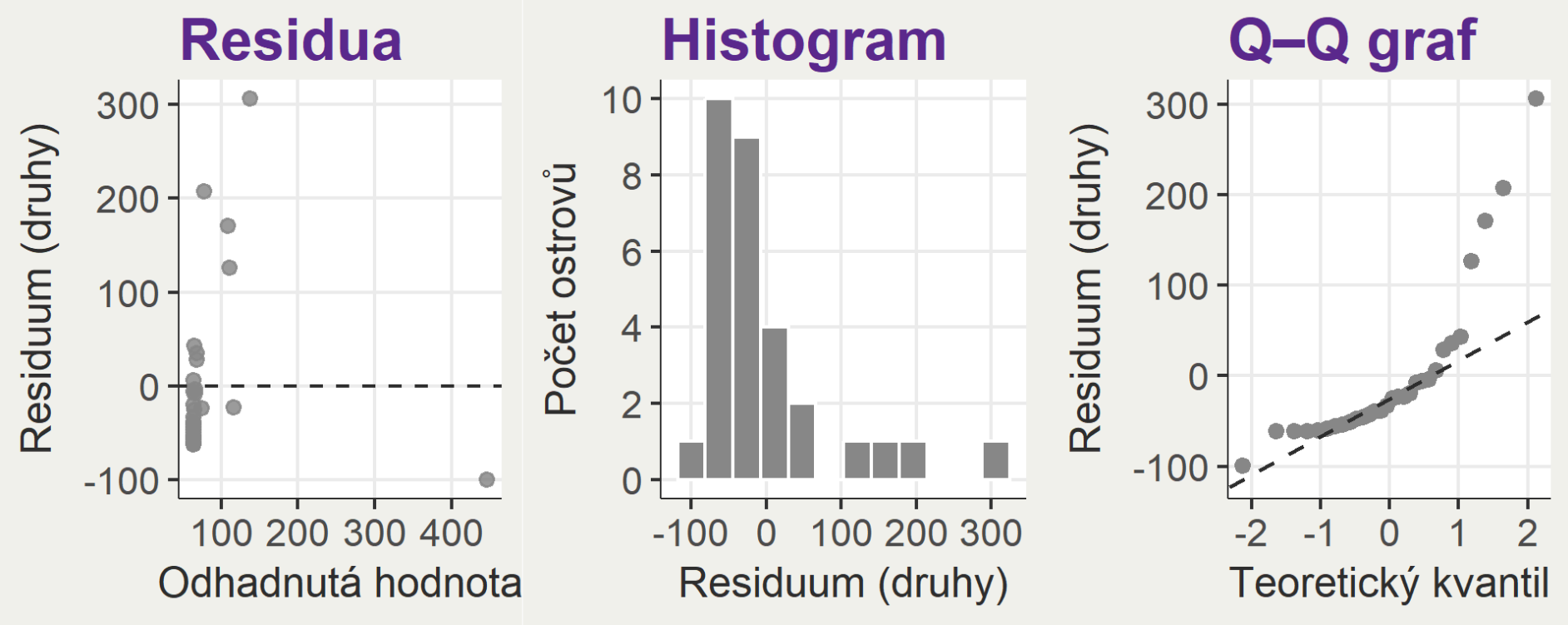
Ve dvojici: Kdyby přímka zachytila všechno systematické, co by v residuích zbylo? Načrtněte jejich histogram.



simulovaná data

Jen náhodné odchylky: souměrný zvon kolem nuly. Takový tvar residuí lineární model předpokládá: přibližně normální rozdělení residuí.

Co ukazují residua přímky?



Ve dvojici: Porovnejte se zvonem z minulého snímku. Najděte alespoň dva vzory, které by u dobře padnoucí přímky neměly být vidět.

Residua tvoří ohyb · jejich rozptyl roste doprava · histogram i Q–Q graf ukazují dlouhý ocas vpravo.

Co musí mít u lineárního modelu přibližně normální rozdělení?

A
odezva (počet druhů)

B
prediktor (plocha ostrova)

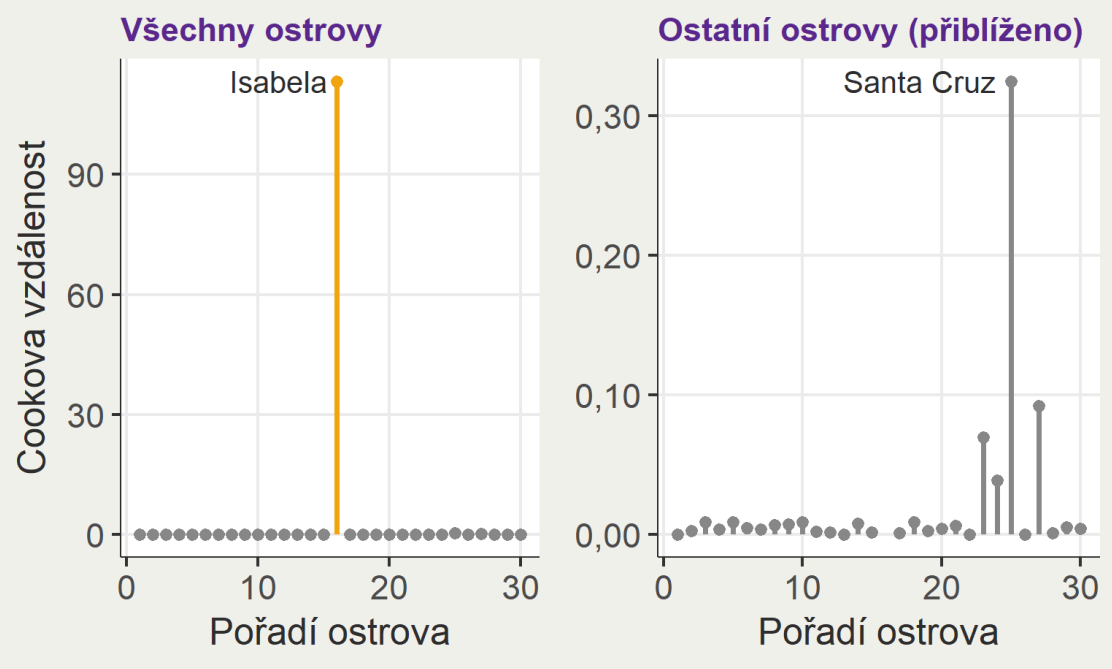
C
residua modelu

D
odezva i prediktor zároveň

Přesvědčte souseda, proč jste vybrali svoji možnost.

Správně je C: model = přímka + náhodné odchylky kolem ní. Předpoklad se týká odchylek, tedy residuí. Šikmá odezva ani šikmý prediktor samy o sobě nevadí.

Jak jeden velký ostrov ovlivní celou přímku?



Cookova vzdálenost: jak moc se přímka změní, když ostrov vynecháme.

Isabela: 113,3 · druhý nejvyšší, Santa Cruz: 0,32

ŠIKMÝ PREDIKTOR

předpoklad modelu neporušuje — o rozdělení prediktoru model nic netvrdí.

PŘESTO PRAKTICKÝ PROBLÉM

jeden ostrov natáčí celou přímku jako dlouhá páka.

Co už o ostrovech víme?

Jak rychle roste počet druhů rostlin s plochou ostrova?

UŽ UMÍME ŘÍCT

- počet druhů i plocha jsou pravostranně šikmé;
- přímka míjí většinu malých ostrovů;
- residua se ohýbají, rozšiřují a mají dlouhý ocas;
- přímku natáčí jediný ostrov.

ZATÍM NEVÍME

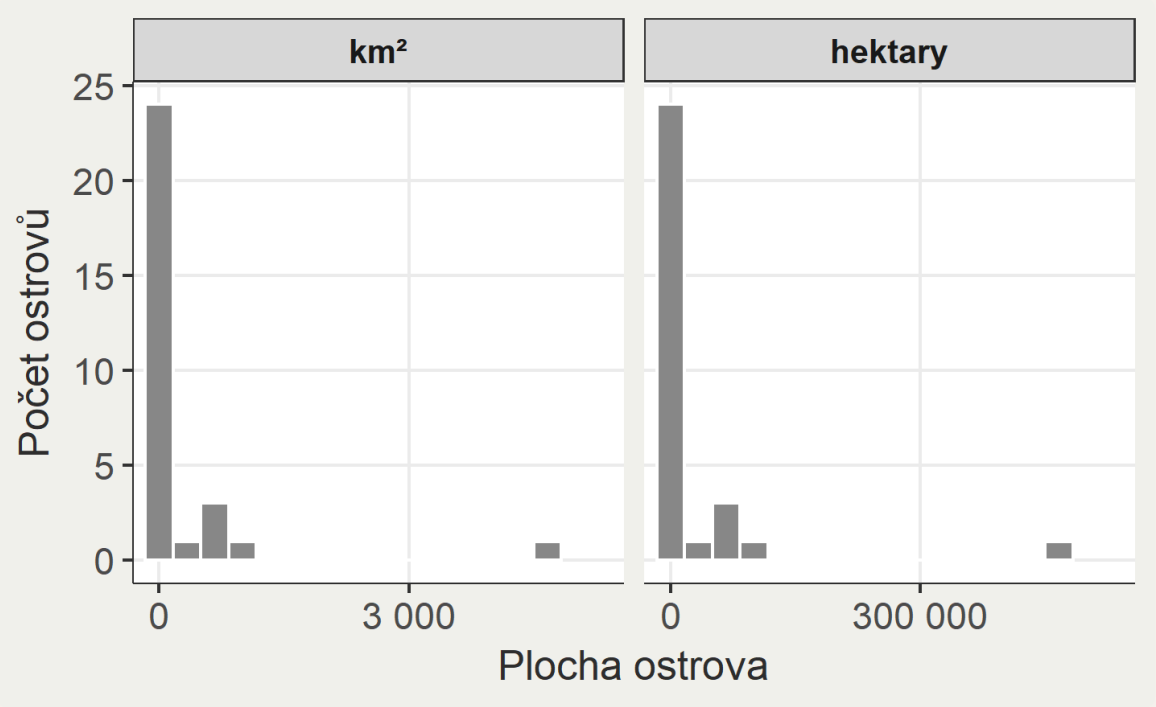
- Jak popsat ostrovy od 0,01 do 4 669 km² jedním modelem?
- Pomůže, když proměnné nějak upravíme? A jakou úpravou?

Transformace: měřítko, nebo tvar?

Změní se tvar rozdělení, když plochu převedeme na hektary?

```
1  data_ostrovy$plocha_ha <-
2    data_ostrovy$plocha_km2 * 100
```

Jeden km² má 100 hektarů.



mění jen měřítko — pořadí ostrovů i poměry vzdáleností mezi nimi zůstávají.

Co se změní, když od plochy odečteme průměr?

```r
data_ostrovy$plocha_centrovana <-
  data_ostrovy$plocha_km2 -
  mean(x = data_ostrovy$plocha_km2)

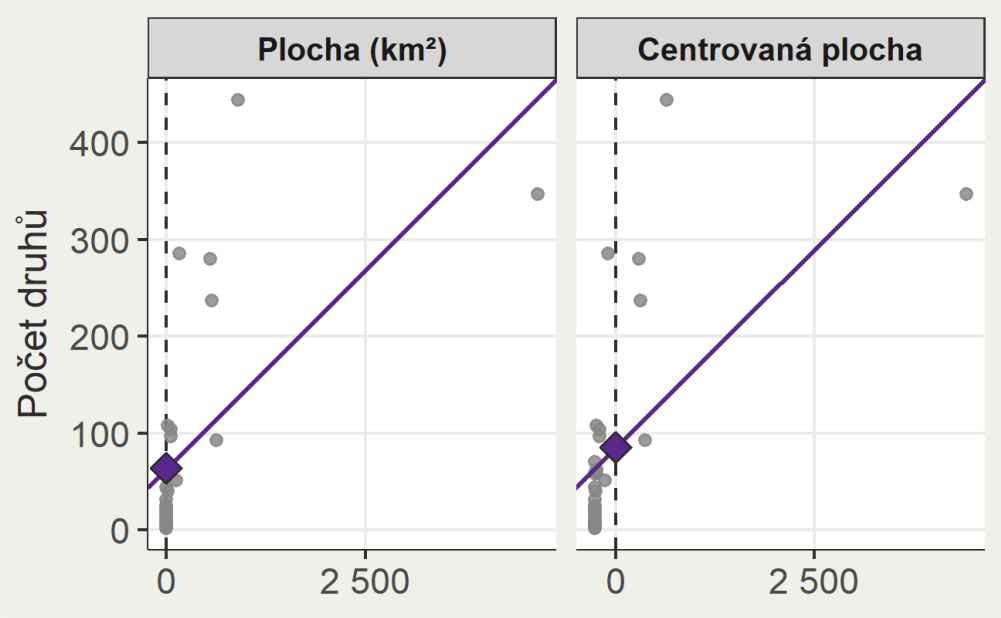
mod_centrovany <- lm(
  formula = pocet_druhu ~ plocha_centrovana,
  data = data_ostrovy
)

coef(object = mod_primka)
```

```
 (Intercept)  plocha_km2
63.78286147  0.08196317
```

```r
coef(object = mod_centrovany)
```

```
       (Intercept) plocha_centrovana
       85.23333333         0.08196317
```



Ve dvojici: Co znamená intercept po centrování?

Sklon se nezměnil (0,082). Intercept 63,8 → 85,2: odhad druhů pro ostrov průměrné plochy, ne pro neexistující ostrov s nulovou plochou. · mění jen měřítko

Jak daleko od průměru leží Fernandina?

Nejvyšší bod Fernandiny **1 494 m** · průměr **424,5 m** · SD **453,0 m**

Spočítejte: o kolik směrodatných odchylek leží Fernandina nad průměrem?

$$\text{z-skóre} = \frac{\text{hodnota} - \text{průměr}}{\text{SD}} = \frac{1494 - 424{,}5}{453{,}0} = 2{,}36$$

```
1  data_ostrovy$vyska_z <- as.numeric(
2    x = scale(x = data_ostrovy$nejvyssi_bod_m)
3  )
```

Tomu se říká **z-skóre** (*z-score*). Z-skóre celé proměnné mají **průměr 0 a SD 1**; kladné znamená nad průměrem.

Je Fernandina výjimečnější plochou, nebo výškou?

z-skóre plochy: 0,43 · z-skóre výšky: 2,36

A · výjimečnější výškou

B · výjimečnější plochou

C · v obou

Z-skóre ukazují na A. Pořadí mezi ostrovy ale říká: plochou 3. největší, výškou 2. nejvyšší — výjimečná je v obou ohledech.

Z-skóre u šikmé veličiny klame. Obrovská Isabela nafoukla SD plochy, takže i 3. největší ostrov vypadá skoro průměrně.

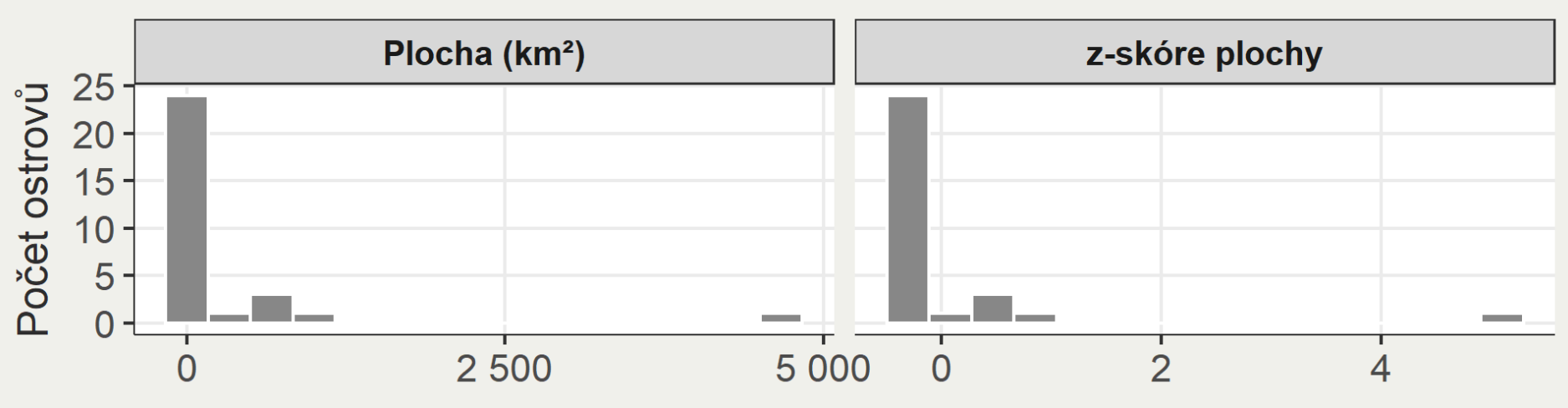
Odstraní z-skóre šikmost?

Plocha ostrovů je extrémně pravostranně šikmá. Bude histogram z-skóre plochy souměrnější?

A · ano

B · trochu

C · ne



Jednotky, centrování i z-skóre jsou **lineární transformace**: násobí a posouvají o stejné číslo. Tvar nezmění.

Jak logaritmus převádí násobky na stejné kroky?

Plocha (km²)	`log(plocha)`
1	0,000
2	0,693
4	1,386
8	2,079
16	? 2,773

Odhadněte poslední řádek bez kalkulačky.

Každé zdvojnásobení přidá stejný krok 0,693. Logaritmus převádí stejné poměry na stejné rozdíly.

```
log(x = c(1, 2, 4, 8))
```

```
[1] 0.0000000 0.6931472 1.3862944 2.0794415
```

```
exp(x = log(x = 8))
```

```
[1] 8
```

```
log(x = 0)
```

```
[1] -Inf
```

`exp()` vrátí původní hodnotu. `log(0)` není konečné číslo: proměnné s nulami logaritmovat nelze.

Které histogramy plochy změní tvar?

Stejnou plochu ostrovů zobrazíme ve čtyřech měřítkách. Které histogramy budou vypadat jinak než histogram v km²?

km²
`plocha_km2`

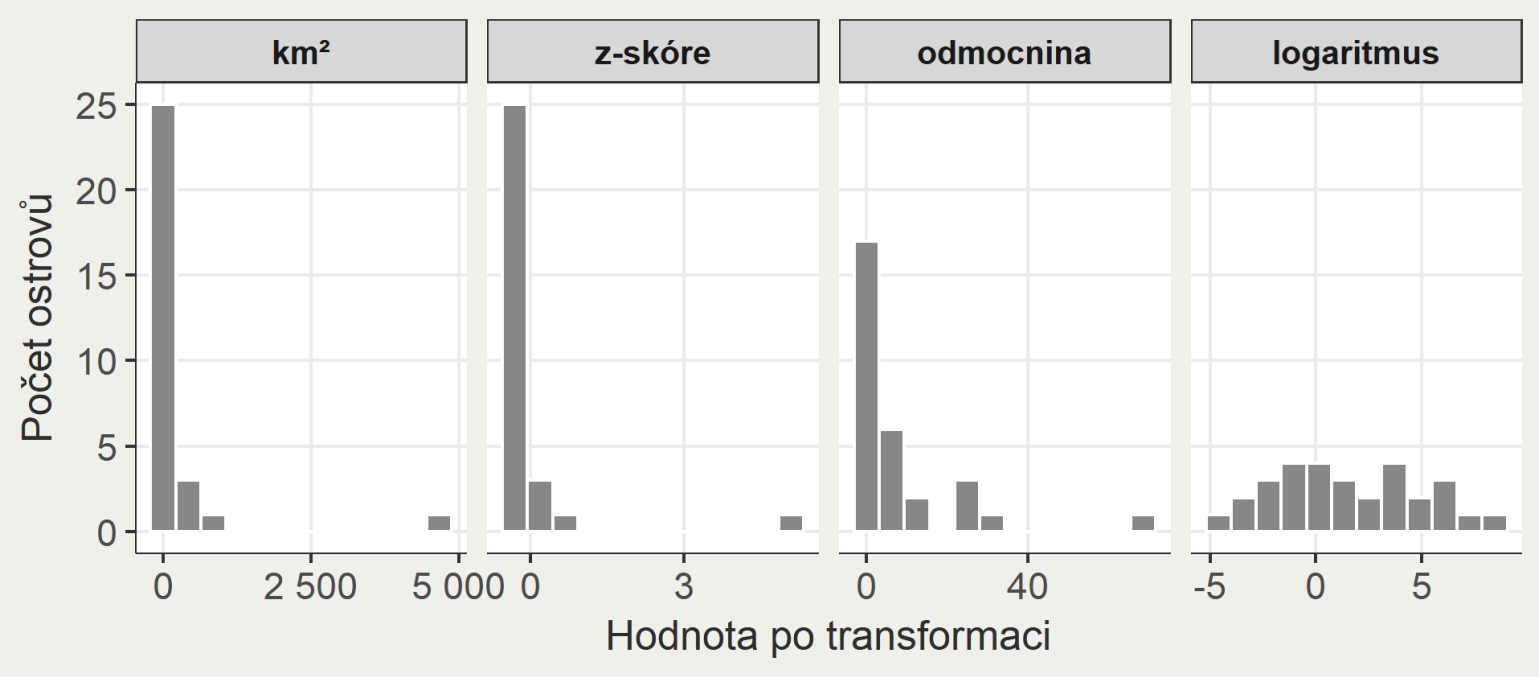
z-skóre
`scale(plocha_km2)`

odmocnina
`sqrt(plocha_km2)`

logaritmus
`log(plocha_km2)`

Hlasujte pro každou kartu zvlášť: změní tvar, nebo ne?

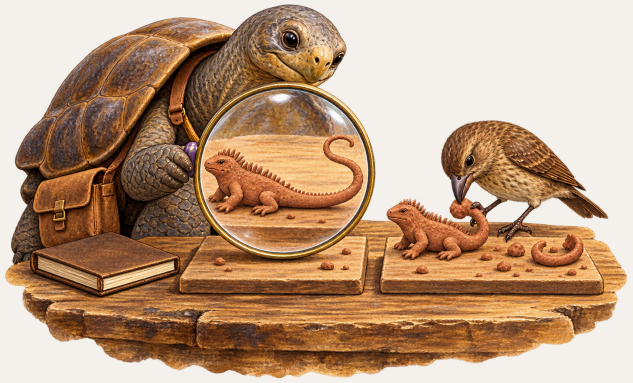
Stejné ostrovy ve čtyřech měřítkách



- z-skóre: stejný tvar jako km²;
- odmocnina: šikmost zmírní, ale neodstraní;
- logaritmus: z extrémně šikmé plochy udělá přibližně souměrné rozdělení. Takové veličině říkáme **log-normální**.

Které transformace mění jen měřítko a které i tvar?

Transformace	V R	Mění	Kdy se hodí
jednotky	`x * 100`	měřítko	různé zdroje dat
centrování	`x - mean(x)`	měřítko	smysluplný intercept
z-skóre	`scale(x)`	měřítko	různé jednotky
odmocnina	`sqrt(x)`	tvar	mírně šikmé počty
logaritmus	`log(x)`	tvar	násobné změny



Ilustrace vytvořená pomocí AI.

Dlouhý ocas zkrátí jen transformace, která mění tvar.

Kde transformace v biologii pomáhají?

LOGARITMUS · KONCENTRACE



Hormony nebo toxiny se mezi vzorky liší o řády.

LOGARITMUS · POPULACE



Populace rostou o procenta, ne o stejný počet.

LOGARITMUS · VELIKOST TĚLA



Hmotnosti savců od myši po slona se liší o řády.

ODMOCNINA · POČTY S NULAMI



Hmyz v pastech: mírná šikmost a prázdné pasti.

Z-SKÓRE · GRADIENT PROSTŘEDÍ



Teplota, srážky a výška v různých jednotkách.

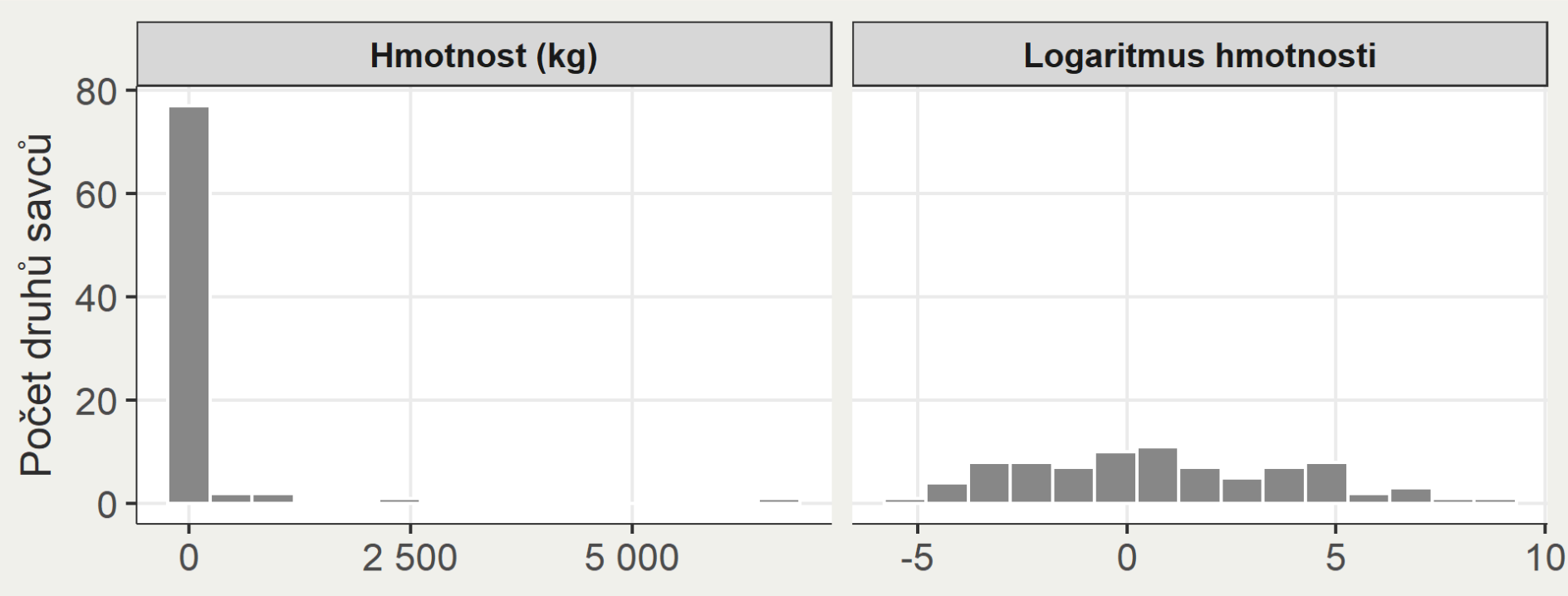
CENTROVÁNÍ · TREND V ČASE



Intercept pro první rok sledování, ne pro rok 0.

Ilustrace vytvořené pomocí AI.

Stejní savci v kilogramech a po logaritmu



Ve dvojici: V první lekci krabicový graf označil 11 druhů jako možná odlehlé. Kolik jich bude po logaritmu?

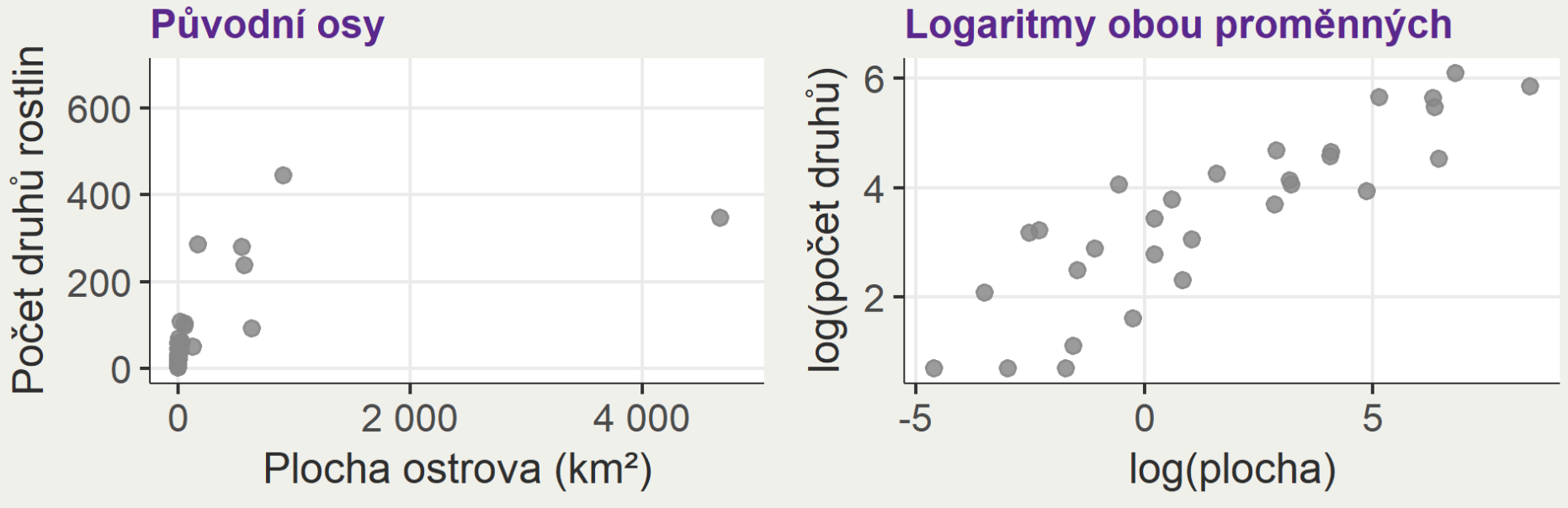
Po logaritmu 0. Největší savci leželi jen v dlouhém ocasu šikmého rozdělení — co je odlehlé, záleží na škále.

Data: `ggplot2::msleep` (83 druhů savců), stejná jako v první lekci

Tvar vztahu

Jak vypadá vztah plochy a počtu druhů na logaritmických osách?

Logaritmus z šikmé plochy udělal souměrné rozdělení. Co udělá s ohybem ve vztahu plochy a počtu druhů?



Ve dvojici: Stejný krok na logaritmické ose = stejný násobek. Co by přímka na pravých osách znamenala pro ostrovy?

Jaké tvary vztahu dávají biologický smysl?

Ve dvojici: Které tvrzení zní nejméně věrohodně pro ostrov o 0,01 km² i pro ostrov o 4 669 km²?

A · PŘÍMKA

Každý další km² přidá stejný počet druhů.

`pocet_druhu ~ plocha_km2`

B · LOGARITMUS PLOCHY

Každé zdvojnásobení plochy přidá stejný počet druhů.

`pocet_druhu ~ log(plocha_km2)`

C · MOCNINNÝ VZTAH

Každé zdvojnásobení plochy vynásobí počet druhů stejným číslem.

`log(pocet_druhu) ~ log(plocha_km2)` · *power law*

D · KVADRATICKÝ ČLEN

Počet druhů roste stále pomaleji, případně se pak obrátí.

`pocet_druhu ~ plocha_km2 + I(plocha_km2^2)`

E · LOGARITMUS ODEZVY

Každý další km² zvýší počet druhů o stejné procento.

`log(pocet_druhu) ~ plocha_km2`

Sadu kandidátů určíme předem — z biologie, ne z výsledků.

Všech pět kandidátů odhadne funkce `lm()`

```r
mod_a_primka <- lm(
  formula = pocet_druhu ~ plocha_km2,
  data = data_ostrovy
)

mod_b_log_plochy <- lm(
  formula = pocet_druhu ~ log(plocha_km2),
  data = data_ostrovy
)

mod_c_mocninny <- lm(
  formula = log(pocet_druhu) ~ log(plocha_km2),
  data = data_ostrovy
)
```

```r
mod_d_kvadraticky <- lm(
  formula = pocet_druhu ~ plocha_km2 +
    I(plocha_km2^2),
  data = data_ostrovy
)

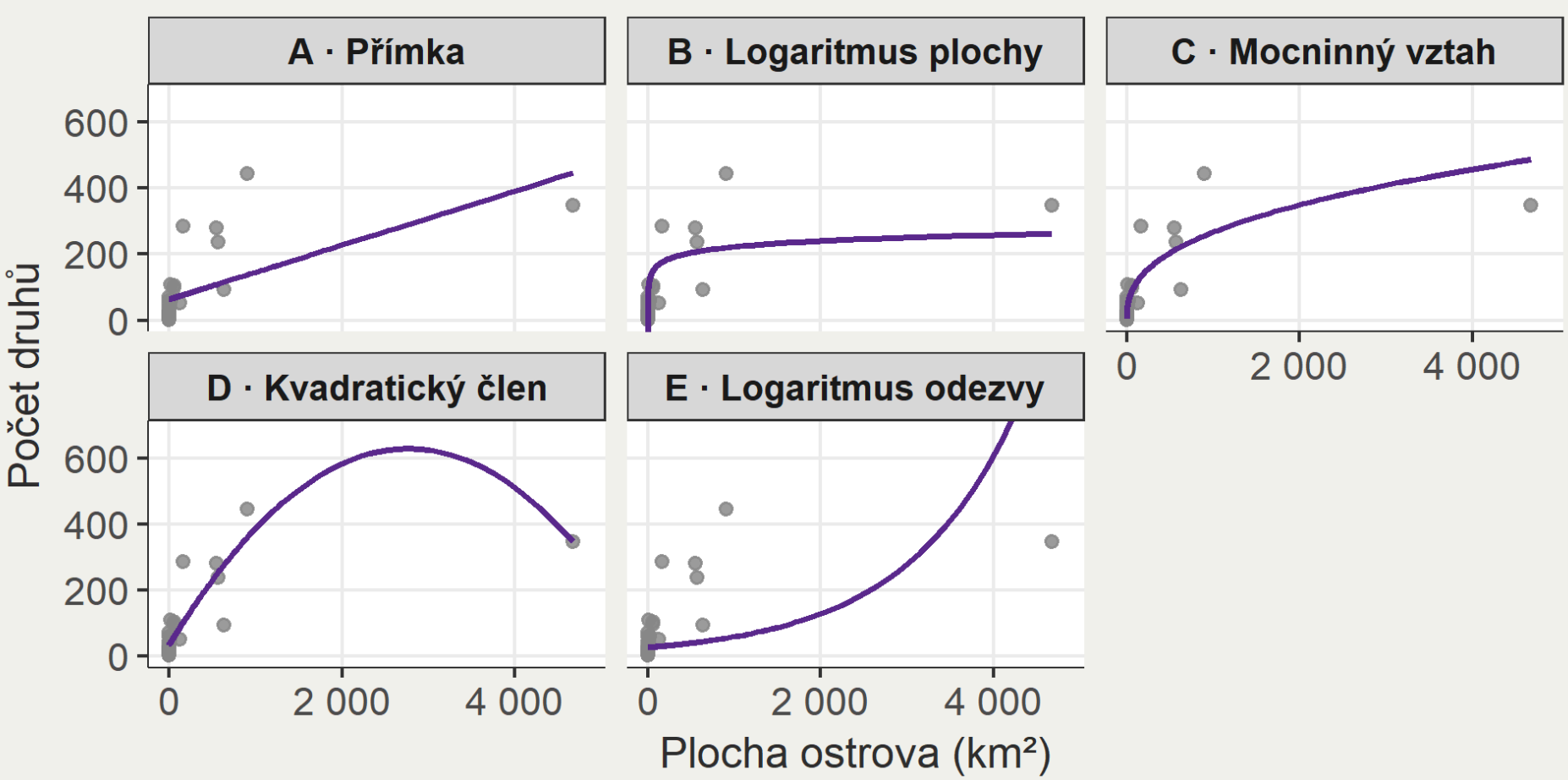
mod_e_log_odezvy <- lm(
  formula = log(pocet_druhu) ~ plocha_km2,
  data = data_ostrovy
)
```

`I()` říká R: spočítej `plocha_km2^2` jako obyčejnou druhou mocninu.

> „Lineární“ model znamená **lineární v koeficientech**, ne nutně přímku v grafu.

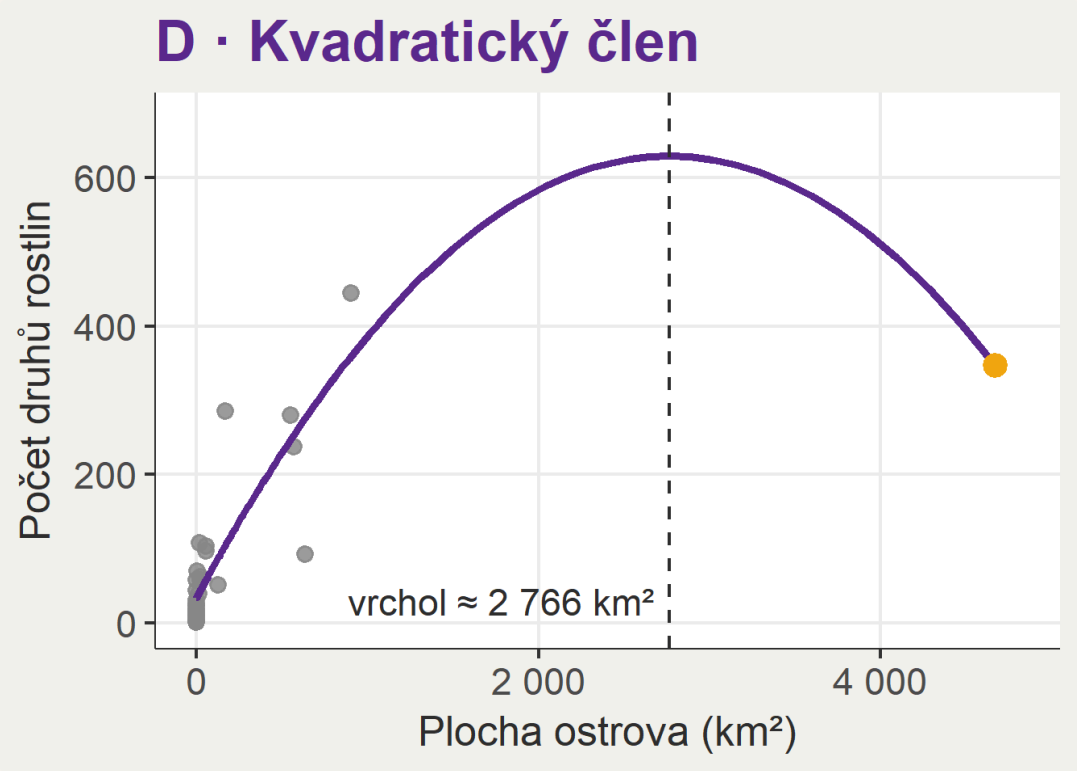
Počet druhů je počet. Modely, které s počty pracují přímo, poznáme v jedné z dalších lekcí.

Jak se kandidátní křivky shodují s ostrovy?



Ve dvojici: Která křivka tvrdí o ostrovech něco, co by biolog nepřijal?

Proč by měl větší ostrov mít méně druhů?





Ilustrace vytvořená pomocí AI.

Za vrcholem leží jen **Isabela** — s Cookovou vzdáleností **142**.

> Parabola se **musí obrátit**. K Isabele se ohýbá jen proto, že ji Isabela táhne.

Je mocninný vztah C podle AIC mnohem lepší než kvadratický člen D?

D · odezva `pocet_druhu` · AIC 334,5
C · odezva `log(pocet_druhu)` · AIC 74,5

A
Ano, AIC modelu C je o 260 nižší, takže C popisuje ostrovy mnohem lépe.

B
Ne, každý z modelů předpovídá jinou odezvu, takže jejich AIC porovnat nelze.

C
Ano, protože mocninný vztah má méně koeficientů, a proto nižší AIC.

D
Ne, protože AIC nelze porovnat mezi žádnými dvěma modely, ani se stejnou odezvou.

Přesvědčte souseda, proč jste vybrali svoji možnost.

Pravidlo z minulé lekce: AIC porovnáváme jen u modelů se stejnou odezvou na stejných pozorováních. Logaritmus stlačí velké hodnoty, takže residua C jsou čísla v úplně jiném měřítku.

Které modely smíme porovnat pomocí AIC?

ODEZVA: POČET DRUHŮ

Kandidát	AIC
A · Přímka	360,2
B · Logaritmus plochy	345,9
D · Kvadratický člen	334,5

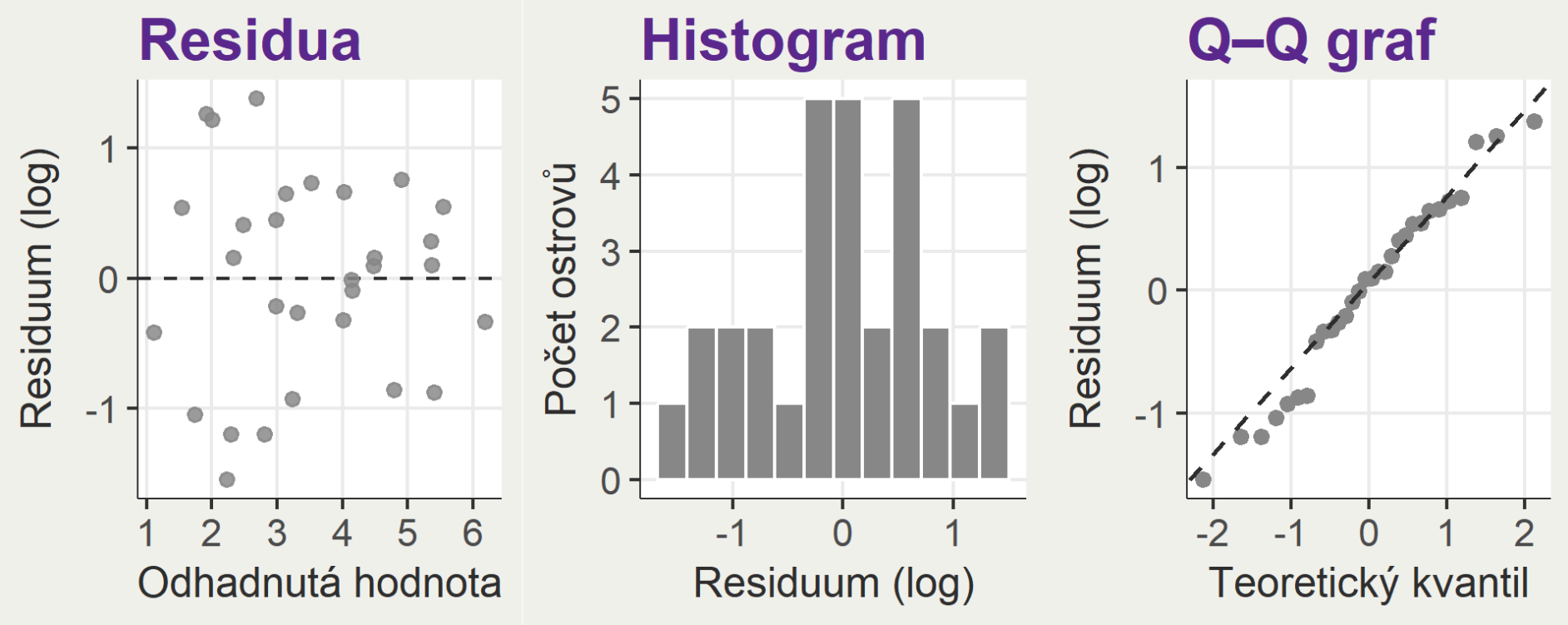
ODEZVA: LOGARITMUS POČTU DRUHŮ

Kandidát	AIC
E · Logaritmus odezvy	110,8
C · Mocninný vztah	74,5

Mezi modely pro počet druhů má nejnižší AIC kvadratický člen D — křivka, kterou jsme odmítli kvůli jejímu tvaru. Nejnižší AIC neznamená biologicky obhajitelný model; rozhoduje i biologie a residua.

A co R^2? Ani to tabulky nespojí: R^2 modelu D je podíl variability počtu druhů, R^2 modelu C podíl variability jeho logaritmu. Napříč škálami porovnáváme křivky na původních osách a vzory v residuích (ne jejich velikost).

Pomohl logaritmus také residuím?

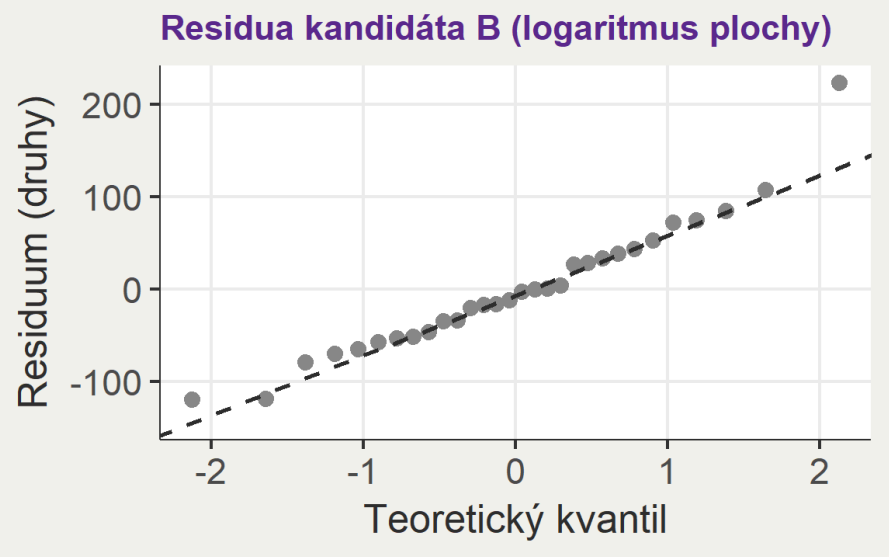


Ve dvojici: Co se změnilo proti residuím přímky?

Pás bez ohybu i rostoucího rozptylu · Q–Q graf blízko přímky · největší Cookova vzdálenost 0,14 místo 113,3. S 30 ostrovy čteme opatrně, ale všechny grafy ukazují stejným směrem.

Může být odezva šikmá, a residua přesto v pořádku?





Odezva má dlouhý ocas vpravo.

Residua B leží blízko přímky.

> Šikmost odezvy z velké části vysvětlil šikmý prediktor. Předpoklad se týká residuí, ne samotné odezvy.

Interpretace po transformaci

Kolikrát víc druhů má dvakrát větší ostrov?

Člen	Odhad	95% interval
intercept	2,904	2,58 až 3,23
log(plocha)	0,389	0,30 až 0,47

```
log(pocet_druhu) ~ log(plocha_km2)
```

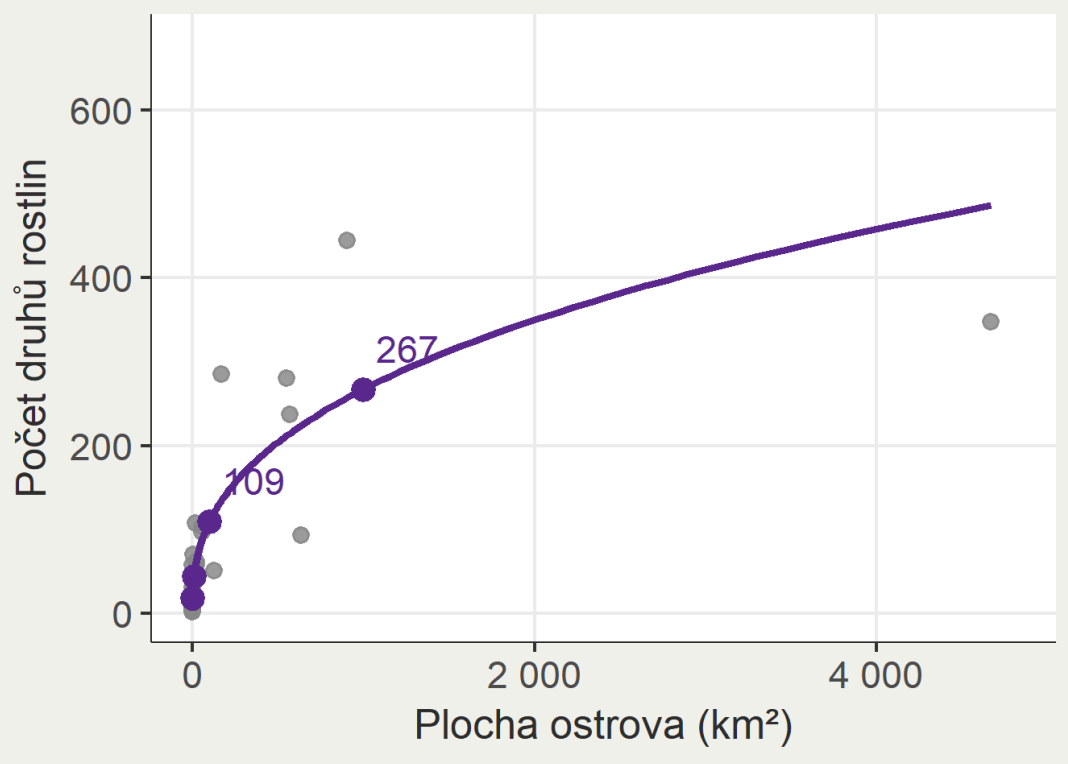
$$\text{násobek druhů} = \text{násobek plochy}^{\text{sklon}}$$

$$2^{0{,}389} = 1{,}31$$

Dvakrát větší ostrov má typicky asi **1,31krát** víc druhů — ne dvakrát.

Ve dvojici: Kolikrát víc druhů má podle modelu **desetkrát** větší ostrov?

Desetkrát větší ostrov má asi 2,45krát víc druhů



$$10^{0{,}389} = 2{,}45$$

95% interval: 2,01 až 2,98krát

Porovnejte se svým odhadem ze začátku přednášky.

Zpětný převod `exp()` dává typický ostrov dané plochy (blízko mediánu), ne průměrný.

Jak číst koeficient, když logaritmujeme jen jednu stranu?

Ve dvojici: Přiřaďte každé větě kandidáta B, D, nebo E.

1 · Každé zdvojnásobení plochy přidá asi 17,8 druhu.
→ B · `pocet_druhu ~ log(plocha_km2)`

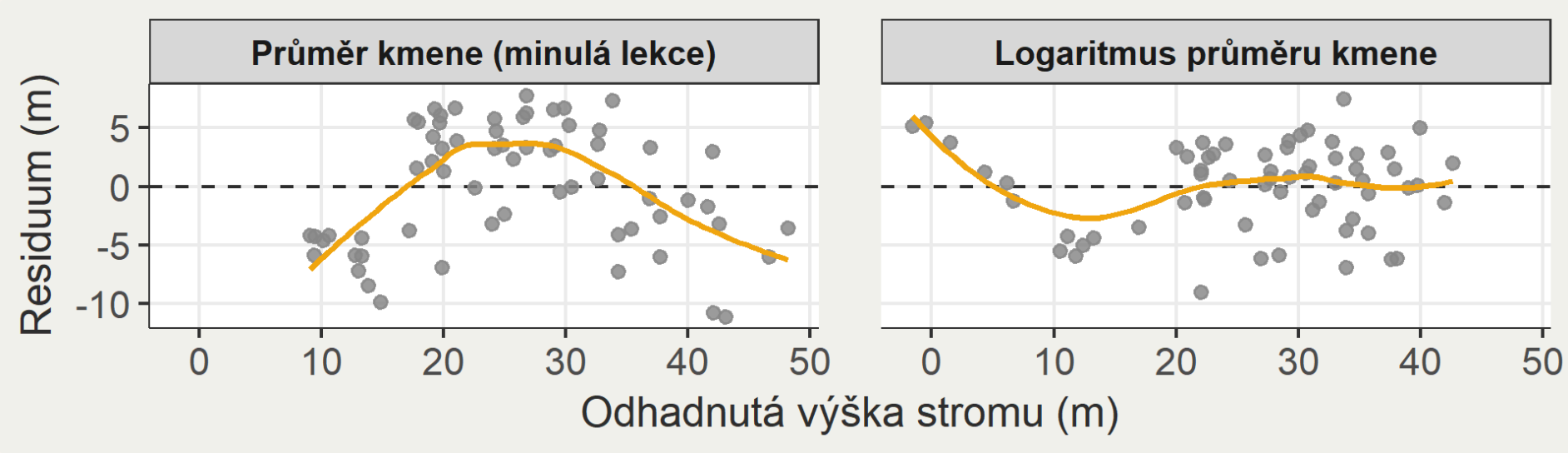
2 · Jak rychle počet druhů roste, se mění podél křivky.
→ D · kvadratický člen: čteme předpovědi, ne jednotlivé koeficienty

3 · Každých 100 km² zvýší počet druhů asi o 8 %.
→ E · `log(pocet_druhu) ~ plocha_km2`

Logaritmus prediktoru → přírůstek na zdvojnásobení · logaritmus odezvy → procento na jednotku · oba → násobek

Zmizí oblouk v residuích stromů po logaritmu průměru kmene?

Výška roste s průměrem kmene **stále pomaleji**. Co udělá s obloukem z minulé lekce logaritmus průměru kmene?



Oblouk se **zploštil**, ale nezmizel: stromy s odhadnutou výškou kolem 10–20 m model stále **přeceňuje**. AIC 395 → **351** — porovnat smíme, odezva je stejná.

Závěr

Odpovězte na úvodní otázku dvěma větami

CO O OSTROVECH VÍME

- 30 ostrovů jednoho souostroví; odezva je počet druhů
- na logaritmických osách leží ostrovy podél přímky
- mocninný vztah: sklon **0,39** (95% interval 0,30 až 0,47)
- dvojnásobná plocha → **1,31×** druhů; desetinásobná → **2,45×**
- residua mocninného vztahu bez vzoru; přímka a parabola neobstály

40 s sami · ve dvojici porovnejte

„Počet druhů roste s plochou ostrova ____, protože ____.“

„Omezením tohoto závěru je ____.“

Použijte aspoň **jedno číslo s interpretací** a **jedno omezení**.

Jak rychle roste počet druhů s plochou ostrova?

Podle **mocninného vztahu** s exponentem asi 0,39 (95% interval 0,30 až 0,47): dvakrát větší ostrov má typicky asi **1,31krát** více druhů, desetkrát větší asi **2,45krát** více.

OPORA

- biologicky zdůvodněný tvar vztahu;
- přímka na logaritmických osách;
- residua bez systematického vzoru.

OMEZENÍ

- jen 30 ostrovů;
- odezva je **počet** — modely pro počty poznáme v jedné z dalších lekcí;
- ostrovy jednoho souostroví nemusí být **nezávislé**.

Kdy transformaci použít a kdy ne?

1. Položte biologickou otázku a prohlédněte tvar rozdělení.
2. Odhadněte nejjednodušší model a zkontrolujte residua.
3. Ptejte se, jaký tvar vztahu předpovídá biologie.
4. Převeďte tvrzení na kandidátní formule s `log()` nebo `I(x^2)`.
5. Porovnejte je na původní škále, residui a AIC se stejnou odezvou.
6. Interpretujte v původních jednotkách: násobky, zdvojnásobení, procenta.

Transformace není rituál. Volíme ji kvůli biologickému tvrzení o vztahu, ne jen proto, aby residua „vypadala normálně“.

Co si odnést

- Mnoho biologických veličin je pravostranně šikmých: hranice v nule a růst násobením.
- Lineární model předpokládá přibližně normální residua, ne odezvu nebo prediktor.
- Jednotky, centrování a z-skóre mění jen měřítko; logaritmus a odmocnina mění tvar.
- Logaritmus převádí násobky na stejné kroky, a proto z mocninného vztahu udělá přímku.
- AIC porovnáváme jen při stejné odezvě; koeficienty po transformaci čteme jako násobky nebo přírůstky.

Jsou ostrovy jednoho souostroví nezávislá pozorování?



Sousední ostrovy si vyměňují semena, sdílejí podnebí i historii osídlení.

Dává 30 ostrovů opravdu 30 nezávislých kousků informace?

Ilustrace vytvořená pomocí AI.