

Introduktion

Anakondaen (*Eunectes murinus*), også kendt som den grønne anakonda, er én af verdens største og tungeste slanger. Den hører til kvælerslangerne og lever primært i tropiske og fugtige områder i Sydamerika – især i Amazonasflodens vandrige lavlandsområder.

Anakondaen er semi-akvatisk og tilbringer størstedelen af sin tid i vand, hvor dens store masse ikke begrænser dens bevægelighed. I vandet jager den fisk, fugle, kapivarer og endda kaimaner. Den dræber sit bytte ved at kvæle det med sin stærke muskulatur.



Figur 1: Anakondaens habitat

Datasæt

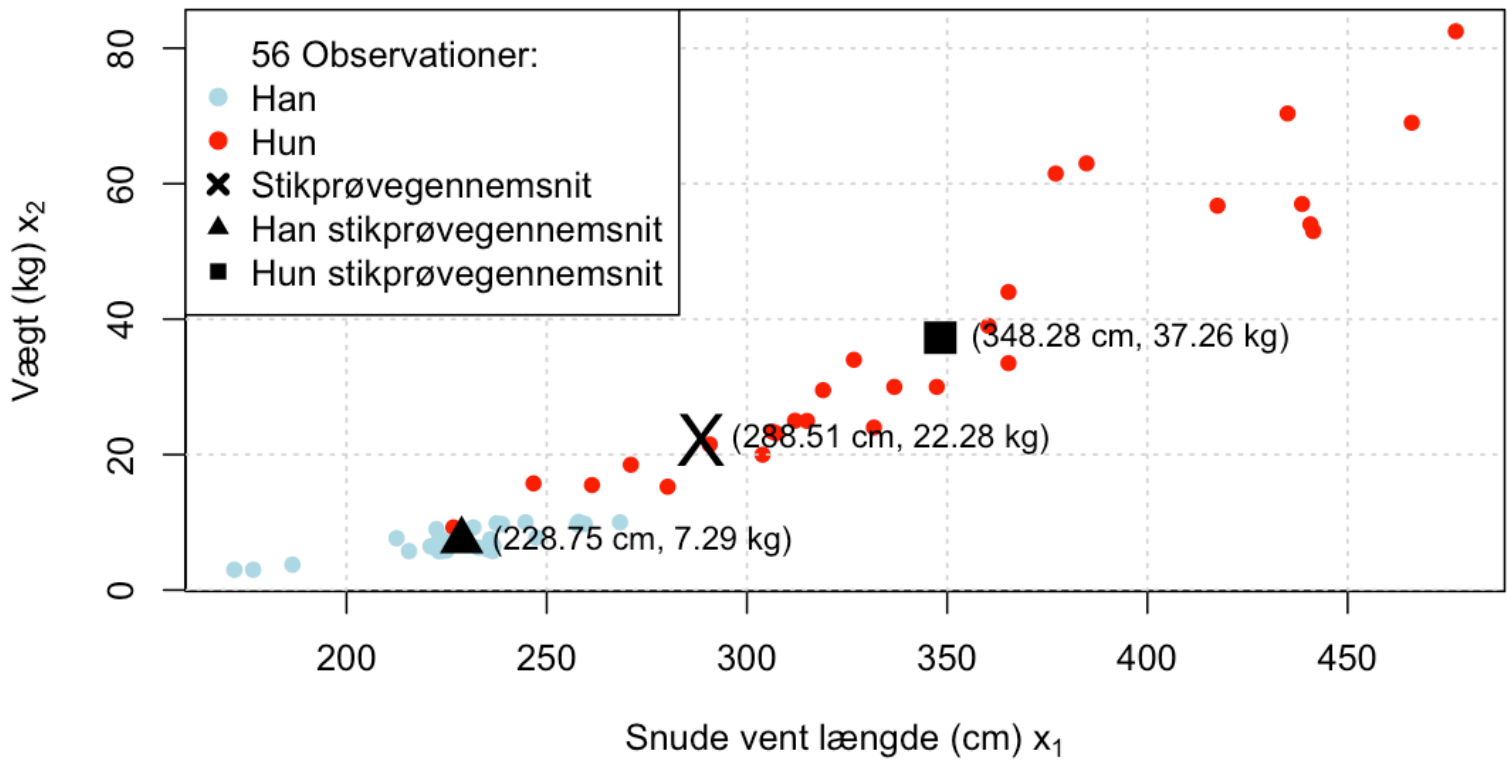
Anakonda-datasættet består af **56 observationer** og **3 variable**:

- **Snude vent længde (cm)** – længden fra snude til den bageste åbning (numerisk)
- **Vægt (kg)** – anakondaens vægt i kilo (numerisk)
- **Køn** - køn som kategorisk variabel med værdierne *Hun* og *Han* (faktor)

Datasættet stammer fra *Table 6.19* i bogen *Applied Multivariate Statistical Analysis* (6. udgave) og udgør grundlaget for vores analyse. Variablerne repræsenterer nøgleegenskaber ved anakondaer og muliggør en direkte sammenligning af længde og vægt mellem kønnene.

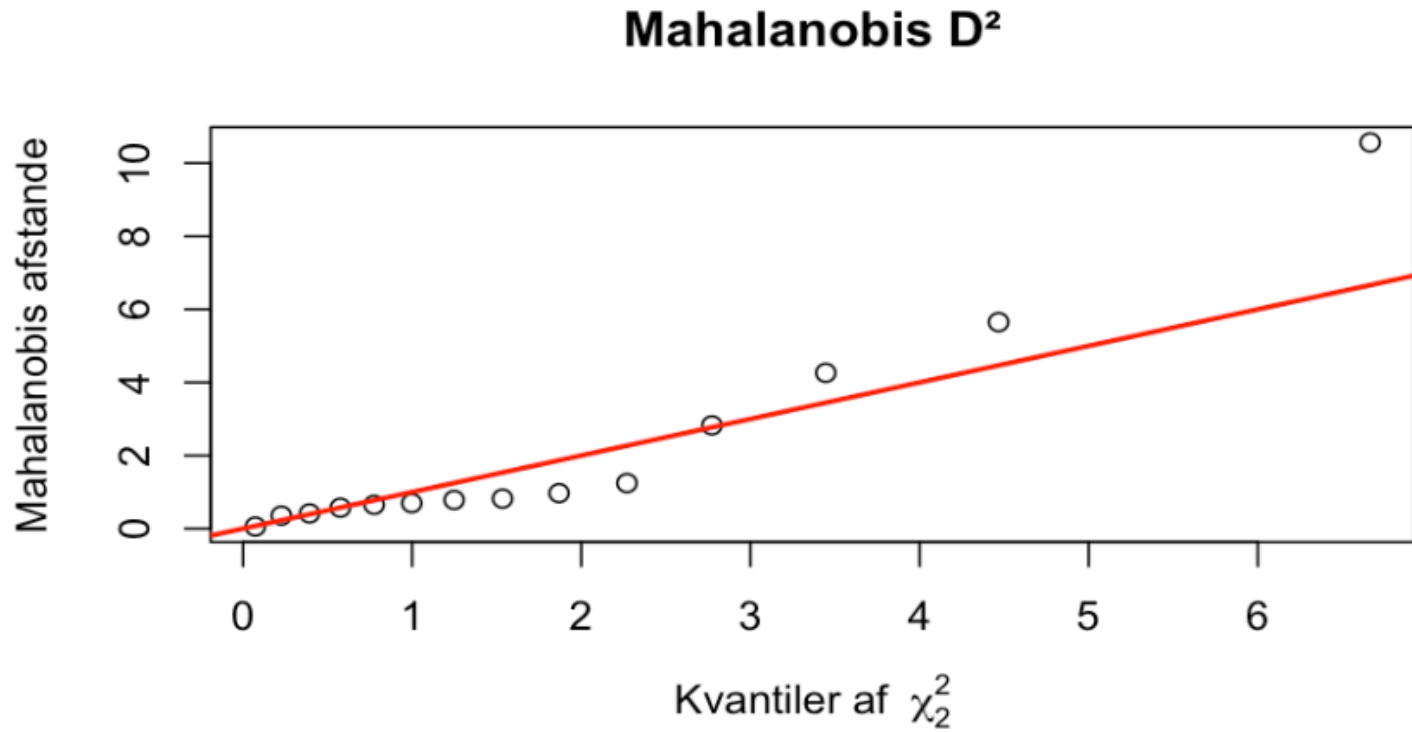
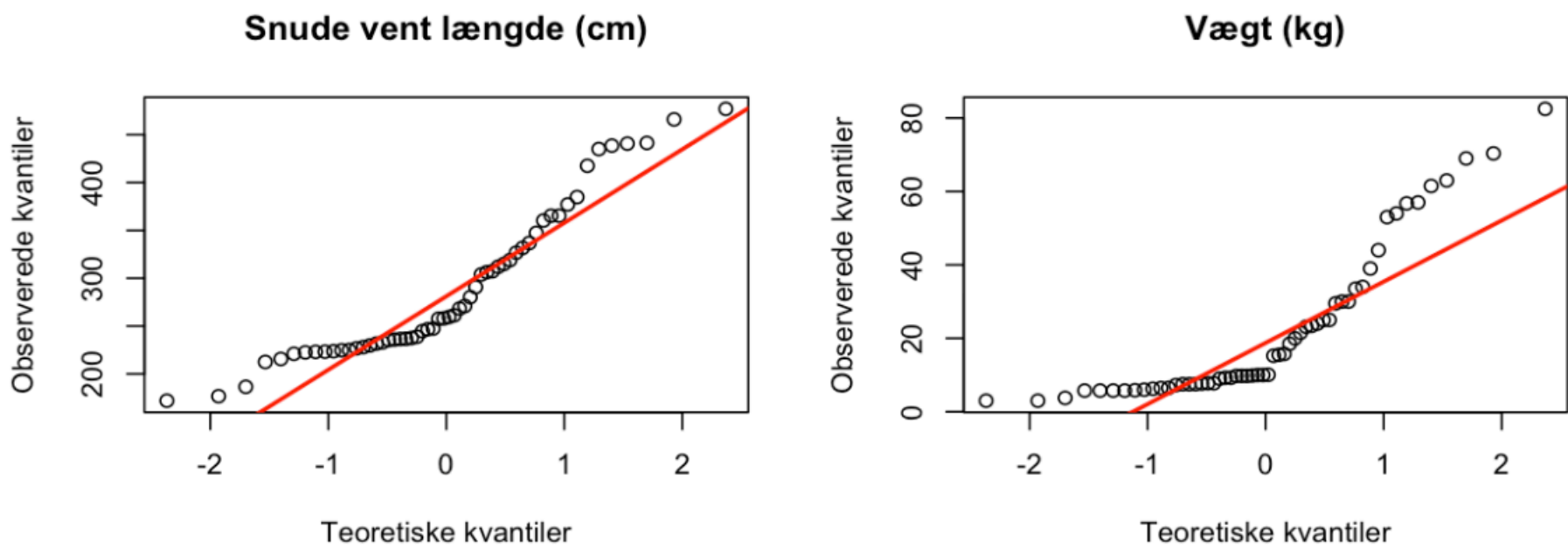
De to numeriske variable indgår som responser i de multivariate analyser, mens køn anvendes som gruppering til at identificere kønsforskelle.

Scatterplot med stikprøvegennemsnit



Figur 2: Scatterplot fordelt på køn

Normalitetstest



Figur 3: Q-Q Plots

De univariate (marginale) og det multivariate Q-Q plot indikerer, at anakonda-dataene ikke følger en normalfordeling.

Vi tester for univariat normalfordelt data, da dette er en forudsætning for at udføre univariat-inferens (T-test). Vi tester for multivariat normalitet, da dette er en forudsætning for at udføre multivariat inferens.

Univariat inferens (T-tests)

Vi har udført to uafhængige T-tests for at undersøge, om der er forskelle i **Snude vent længde (cm)** og **Vægt (kg)** mellem han- og hun-anakondaer.

I de univariate T-tests tester vi om $\mu_{\text{Hun}} = \mu_{\text{Han}}$

- H_0 : Der er ingen forskel i gennemsnitlig Snude vent længde mellem kønnene.
- H_1 : Der er forskel i gennemsnitlig Snude vent længde mellem kønnene.

- H_0 : Der er ingen forskel i gennemsnitlig vægt mellem kønnene.
- H_1 : Der er forskel i gennemsnitlig vægt mellem kønnene.

Variabel	μ_{Hun}	μ_{Han}	T-værdi	P-værdi	95% KI
Snude vent længde (cm)	348.28	228.75	8.76	$< 5.98 \cdot 10^{-12}$	[92.17, 146.88]
Vægt (kg)	37.26	7.29	7.85	$< 1.74 \cdot 10^{-10}$	[22.32, 37.63]

Tabel 1: Resultater fra uafhængige T-tests for længde og vægt mellem kønnene

Begge tests viser ekstremt lave p-værdier ($p \ll 0.05$), hvilket betyder, at vi **forkaster nulhypotesen og konkluderer, at både længde og vægt afhænger af køn**. Hun-anakondaer er i gennemsnit ca. **120 cm længere** og **30 kg tungere** end hanner – en forskel, der både er **statistisk signifikant** og biologisk relevant.

Multivariat inferens

Her tester vi Hotelling's T^2 -test:

$$H_0 : \mu_{\text{Hun}} = \mu_{\text{Han}} \quad \text{mod} \quad H_A : \mu_{\text{Hun}} \neq \mu_{\text{Han}}$$

Denne test undersøger, om der er en samlet forskel i middelvektorerne for **Snude vent længde** og **Vægt** mellem hunner og hanner.

Resultat:	T^2	F	P-værdi
	76.9153	37.7455	$6.4296 \cdot 10^{-11}$

Da p-værdien er ekstremt lav ($p \ll 0.05$), **forkaster vi nulhypotesen** $H_0 : \mu_{\text{Hun}} = \mu_{\text{Han}}$.

MANOVA

Vi anvender en multivariat variansanalyse (MANOVA) til at teste for samlede kønsforskelle i længde og vægt.

Wilks' lambda: 0.4125 F(2, 53): 37.75 P-værdi: $6.43 \cdot 10^{-11}$

Resultatet er signifikant ($p \ll 0.05$) og understøtter tidligere analyser: længde og vægt afhænger stærkt af køn. MANOVA med Bartlett's korrektion giver samme resultat.

Test for kovariansmatricer (Box's M-test)

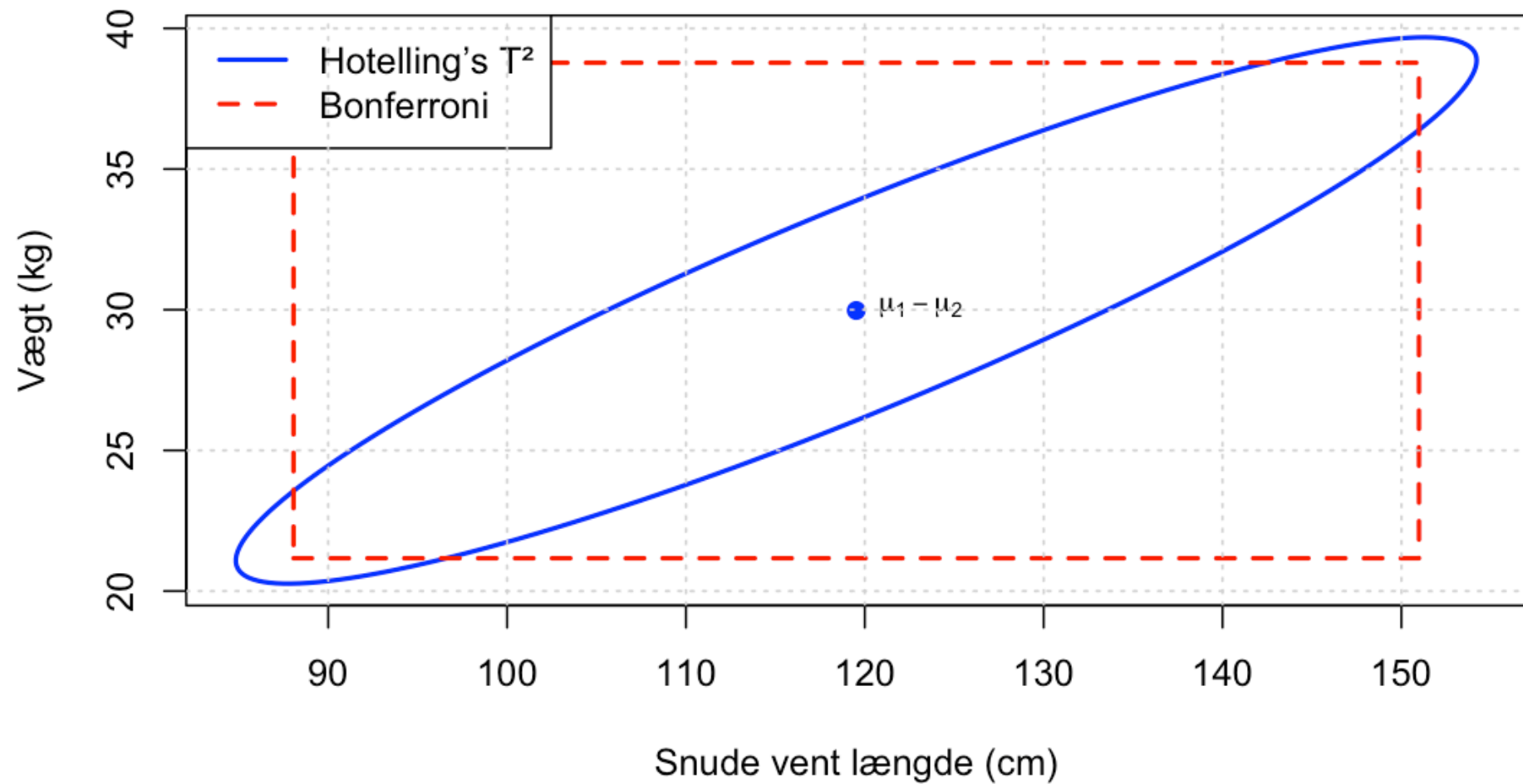
Box's M-test vurderer, om kovariansmatricerne er ens mellem grupperne:

$$H_0 : \Sigma_{\text{Hun}} = \Sigma_{\text{Han}} \quad \text{mod} \quad H_A : \Sigma_{\text{Hun}} \neq \Sigma_{\text{Han}}$$

Resultat:	χ^2	df	p-værdi
	100.325	3	$1.323 \cdot 10^{-21}$

Da p-værdien er ekstremt lav ($p \ll 0.05$), forkaster vi nulhypotesen. Der er signifikant forskel i kovariansstrukturen mellem kønnene.

Simultane konfidensintervaller



Figur 4: Simultane konfidensintervaller ikke indeholdende punktet (0,0)

Snude vent længde (cm): [84.8 ; 154.2], [88.1 ; 151.0] Vægt (kg): [20.3 ; 39.7], [21.2 ; 38.8]

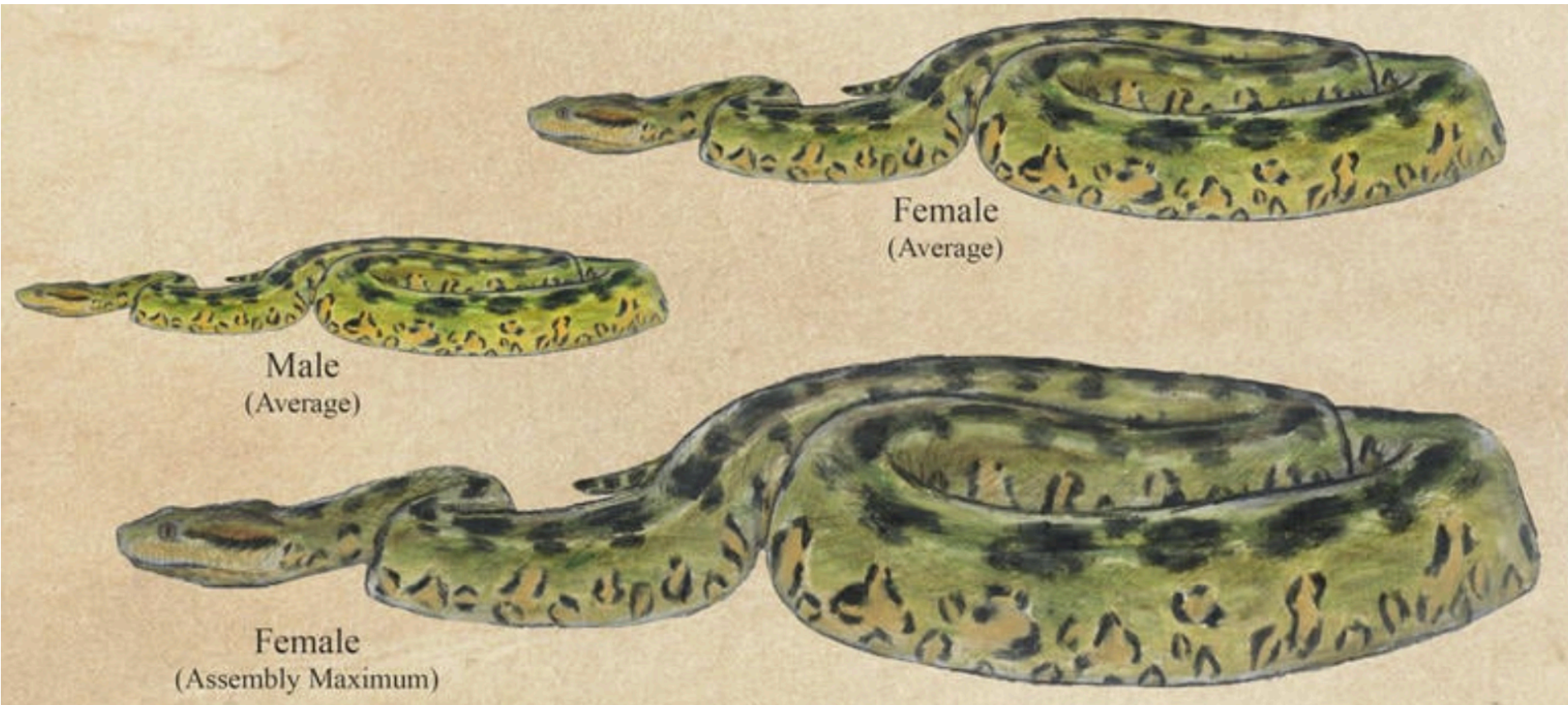
Antagelser og resultater

Resultaterne fra univariat- og multivariat-inferens kræver, at man kritisk forholder sig til de underliggende antagelser. **Univariat T-test bygger på antagelsen om marginal normalfordeling**, imens **Hotellings T2-test teoretisk forudsætter, at data følger en multivariat normalfordeling**. Manova og Box's M-test antager, at **grupperne har ens kovariansstrukturer**. Normalitetsantagelsen er central, særligt i tolkningen af p-værdier fra tests som Box's M, hvor teststørrelsen følger en chi-i-anden-fordeling under antagelse af normalitet. I vores analyser indikerer Q-Q plots dog, at **denne antagelse ikke er opfyldt** – hverken univariat eller multivariat. Derfor bør resultaterne tolkes med forsigtighed, da brud på disse forudsætninger kan påvirke testenenes gyldighed og pålidelighed.

Konklusion

Der er signifikant forskel i både længde og vægt mellem han- og hun-anakondaer.

De univariate og multivariate hypotesetest forkaster klart nulhypotesen om ens middelværdier. Hun-anakondaer er typisk både længere og tungere end hanner – et fænomen kendt som **omvendt kønsdimorfisme**, hvor hunnen er størst. Denne forskel spiller en vigtig rolle under parringsritualer, hvor flere hanner samles omkring en stor hun. I de simultane konfidensintervaller ligger punktet (0, 0) tydeligt udenfor, hvilket bekræfter forskellen mellem kønnene.



Figur 5: Illustration af anakondaens størrelser