

## ODPOWIEDZI DO ZADAŃ

1. Wyświetlenie wymaganej informacji osiągnąć najłatwiej tworząc dodatkowy obiekt zawierający tylko wpisy bez skoku, a następnie sprawdzenie ile ich tam jest. Warto zapamiętać, że dla obiektu typu `dataframe` funkcja `length` zwróci ilość zmiennych (kolumn). Poniższy kod zawiera właściwe rozwiązanie:

```
> gryfy2 <- gryfy[is.na(gryfy$TARSUS),]  
  
> length(gryfy2[,1])
```

```
[1] 401
```

Funkcja `is.na()` jest znacznie lepszym sposobem sprawdzenia czy coś ma wartość NA czy nie (lepszym niż `X==NA`, który nie wie, czemu nie daje zbyt dobrych wyników).

2. Poniżej zamieszczam kolejne etapy budowy modelu statystycznego: definicja priora, budowa modelu uwzględniającego dymorfizm płciowy (czyli ustalony efekt płci) oraz obliczenie odziedziczalności.

```
> var.p=var(gryfy$TARSUS,na.rm=T)  
  
> prior.uni = list(R=list(V=var.p/2,n=1),  
+ G=list(G1=list(V=var.p/2,n=1)))  
  
> modeluni <- MCMCglmm(TARSUS~1+SEX, random=~animal,  
+ prior=prior.uni, data=gryfy, ped=gryfyped, verbose=FALSE)
```

Warning message:

```
In MCMCglmm(TARSUS ~ 1 + SEX, random = ~animal, prior = prior.uni, :  
  some combinations in animal do not exist and 225 missing records  
  have been generated
```

```
> h2.uni <- modeluni$VCV[,1]/(modeluni$VCV[,1]+modeluni$VCV[,2])  
  
> posterior.mode(h2.uni)  
      var1  
0.4051587  
  
> HPDinterval(h2.uni)  
      lower      upper  
var1 0.2669416 0.5920814  
attr(,"Probability")
```

```
[1] 0.95
```

```
> modeluni2 <- MCMCglmm(TARSUS~1+SEX, prior=prior.uni,  
+ data=gryfy, verbose=FALSE)
```

```
> modeluni$DIC
```

```
[1] 4141.892
```

```
> modeluni2$DIC
```

```
[1] 4258.725
```

Ostatnie linijki pokazują formalny sposób selekcji modeli. Odziedziczalność jest istotnie różna od zera i jej przedział ufności leży daleko od wartości zero, należy jednak takie „podejrzenia” potwierdzić selekcją modeli opartą na DIC (która daje identyczną odpowiedź – model zawierający czynnik losowy genetyczny ma niższą wartość DIC).

Pobieżne obejrzenie wykresów serii czasowych dla czynnika losowego oraz residuali wskazuje nad istotną autokorelację (widoczny trend w serii czasowej). Aby upewnić się do co poprawności wyciąganych wniosków modele należałoby uruchomić z dodatkowymi parametrami, np.: `nitt = 150000`, `burnin = 30000`, `thin = 50`.

Uwzględnienie dwóch pozostałych źródeł zmienności wymaga dopisania, odpowiednio, roku urodzenia `BYEAR` oraz czynnika `MOTHER` do specyfikacji efektów losowych. Oczywiście – zmodyfikowania wymaga prior (należy dopisać odpowiednią ilość elementów `G` oraz ich wariancje). Powinniśmy zauważyć 2 istotne rzeczy: uwzględnienie efektu `BYEAR` powoduje nieznaczny spadek wariancji resztkowej (`units`) bez większego wpływu na odziedziczalność. Uwzględnienie efektu `MOTHER` powoduje natomiast istotną zmianę w odziedziczalności (spadek) co pokazuje jak istotne jest uwzględnianie genetycznych efektów matczynych; potrafią one powodować istotne zawyżenie oszacowań odziedziczalności w sytuacji, gdy się ich nie kontroluje.

3. Poniższy kod pokazuje konstrukcję prionów oraz modelu dwuzmiennowego, zawierającego efekty matczyne oraz związane z rokiem urodzenia.

```
> phen.var<-matrix(c(var(gryfy$BWT,na.rm=TRUE),0,0,  
+ var(gryfy$TARSUS,na.rm=TRUE)),2,2)
```

```
> prior.bi<-list(G=list(G1=list(V=phen.var/4,n=2),  
+ G2=list(V=phen.var/4,n=2),  
+ G3=list(V=phen.var/4,n=2)),  
+ R=list(V=phen.var/4,n=2))
```

```

> modelbi<-MCMCglmm(cbind(BWT,TARSUS)~trait-1+trait:SEX,
+ random=~us(trait):animal+us(trait):BYEAR+us(trait):MOTHER,
+ rcov=~us(trait):units,
+ family=c("gaussian","gaussian"),
+ pedigree=gryfyped,data=gryfy,
+ nitt=200000,thin=150,burnin=40000,
+ prior=prior.bi,verbose=FALSE)

> rG <- modelbi$VCV[, "animal.trait.BWT.TARSUS" ] /
+ sqrt(modelbi$VCV[, "animal.trait.BWT.BWT" ] *
+ modelbi$VCV[, "animal.trait.TARSUS.TARSUS" ])

```

Zauważmy, że zamiast odwoływać się do konkretnych kolumn elementów obiektu wynikowego (np. `modelbi$VCV[, 2]`), podajemy ich nazwy (`modelbi$VCV[, "animal.trait.BWT.TARSUS" ]`). Poniżej obliczamy korelację genetyczną skoku i masy urodzeniowej. Nie jest ona równa 1, jest jednak wysoka i różna od zera, co daje nam szansę na prowadzenie skorelowanej selekcji skoku, opartej na masie urodzeniowej.

```

> posterior.mode(rG)
var1
0.7735107

> HPDinterval(rG)
lower upper
var1 0.5699206 0.8782866
attr(,"Probability")
[1] 0.95

```

Korelację matczyną liczy się bardzo podobnie – dzieląc odpowiednią kowariancję (dla efektu matki) przez pierwiastek iloczynu odpowiednich wariancji:

```

> rM<-modelbi$VCV[, "MOTHER.trait.BWT.TARSUS" ] /
+ sqrt(modelbi$VCV[, "MOTHER.trait.BWT.BWT" ] *
+ modelbi$VCV[, "MOTHER.trait.TARSUS.TARSUS" ])

> posterior.mode(rM)
var1
-0.4911696

> HPDinterval(rM,0.95)
lower upper
var1 -0.7359405 -0.1739729
attr(,"Probability")
[1] 0.95

```

Zauważmy, że korelacja matczyna jest istotnie mniejsza od zera. Oznacza to, że efekt ten mógłby potencjalnie przeszkadzać w selekcji opartej na wadze urodzeniowej. Dokładne zbadanie jego wpływu wymagałoby jednak dalszych badań (np. oddzielenia genetycznego efektu matki od

efektu tzw. wspólnego środowiska – dałoby to wskazówki dotyczące sposobów „ominięcia” takiej ujemnej korelacji na poziomie efektów matki).

W przedstawionym modelu zastosowaliśmy strukturę kowariancji typu `us ( )`, powodującą oszacowania wszystkich parametrów macierzy VCV. Gdybyśmy jednak dysponowali danymi tylko dot. długości skoku lub tylko masy dla pojedynczego osobnika, skorelowania wariancji resztkowej nie można już założyć, stąd konieczne byłoby zastosowanie wariancji typu `idh ( )` dla czynnika `units` (argument `rcov` w modelu).

4. Skonstruowanie wymaganej tabeli jest łatwe. Na początek dodajmy do istniejącej tabeli 2 kolumny na nowe dane:

```
> gryfy <- cbind(gryfy, TARSM=NA, TARSF=NA)
```

Następnie przenosimy dane z kolumny TARSUS **różne** od NA i zawierające cyfrę 1 (samiec) w kolumnie SEX do kolumny TARSM, zaś dane zawierające cyfrę 2 w kolumnie SEX do kolumny TARSF. Dla przykładu, pierwszy wiersz mówi: wybierz te wiersze z obiektu gryfy, które w kolumnie SEX zawierają wartość 1, i przypisz odpowiadającą im wartość z kolumny 6 (TARSUS) do kolumny 7 (TARSM). W wyniku dostajemy odpowiednie dane.

```
> gryfy[gryfy$SEX==1,][,6]->gryfy[gryfy$SEX==1,][,7]
> gryfy[gryfy$SEX==2,][,6]->gryfy[gryfy$SEX==2,][,8]
> head(gryfy)
```

	animal	MOTHER	BYEAR	SEX	BWT	TARSUS	TARSM	TARSF
1	1029	1145	968	1	10.77	24.77	24.77	NA
2	1299	811	968	1	9.30	22.46	22.46	NA
3	643	642	970	2	3.98	12.89	NA	12.89
4	1183	1186	970	1	5.39	20.47	20.47	NA
5	1238	1237	970	2	12.12	NA	NA	NA
6	891	895	970	1	NA	NA	NA	NA