

METODY PREDYKCJI WARTOŚCI GENETYCZNEJ W PAKIECIE R

Lista zadań

1. Proszę skonstruować w pakiecie R funkcję o nazwie *heritability*, która dla podanej wariancji genetycznej oraz wariancji fenotypowej oblicza wartość odziedziczalności cechy. Należy uniemożliwić działanie funkcji jeżeli użytkownik poda ujemne wartości wariancji lub wariancja genotypowa będzie przekraczała wartość wariancji fenotypowej.
2. Proszę przekształcić funkcję *heritability* tak, aby dla trzech argumentów (odziedziczalności, wariancji genetycznej, wariancji fenotypowej) jeden mógł pozostać nieznanym i wyliczony przez funkcję, np. w argumentach podajemy wariancję genetyczną oraz odziedziczalność, a funkcja oblicza wariancję fenotypową. Przykładowy wynik działania funkcji powinien być następujący:

```
-----  
> heritability(var_a = 1, h2 = 0.2)  
$odziedziczalność  
[1] 0.2
```

```
$`wariancja genetyczna`  
[1] 1
```

```
$`wariancja fenotypowa`  
[1] 0.2
```

```
-----  
> heritability(var_a = -1, h2 = 0.2)  
Error in heritability(var_a = -1, h2 = 0.2) :  
  wariancja genetyczna nie może być mniejsza od zera!
```

```
-----  
> x = heritability(var_a = 1, h2 = 0.2)  
> x$odziedziczalność  
[1] 0.2  
-----
```

3. Proszę skonstruować w pakiecie R funkcję o nazwie *gen_progress*, która dla danej średniej wartości cechy w pokoleniu rodzicielskim (P_0), średniej wartości cechy osobników wybranych na rodziców następnego pokolenia (P_r), odstępów pokoleń (T), odziedziczalności (h^2), wariancji genetycznej (σ_a^2) oraz wariancji fenotypowej (σ_p^2) oblicza standaryzowaną różnicę selekcyjną, oczekiwany postęp hodowlany na jedno pokolenie i na jeden rok. Przykładowy wynik użycia funkcji:

```
> gen_progress(P0 = 148.8, Pr = 157.49, T = 3, h2 = 0.49, var_a = 122.3236,
var_p = 249.64)
$`standaryzowana różnica selekcyjna`
[1] 0.55

$`postęp hodowlany`
[1] 4.2581

$`postęp hodowlany na 1 rok`
[1] 1.419367
```

4. Proszę przekształcić funkcję *gen_progress* tak, aby wykorzystać w niej funkcję *heritability*.
5. Proszę bez użycia, a następnie z użyciem komputera wykonać następujące obliczenia oraz określić wymiary macierzy wynikowych:

a) $\begin{bmatrix} 1 & 5 & 3 \\ 2 & -3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & -3 & 5 \\ -1 & 4 & -2 \\ 3 & -1 & 1 \end{bmatrix}$

b) $\begin{vmatrix} -1 & 5 & 4 \\ 3 & -2 & 0 \\ -1 & 3 & 6 \end{vmatrix}$

c) $\begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & 0 \\ 1 & -1 & 1 \end{bmatrix}^{-1}$

6. Z wykorzystaniem pakietu R, proszę:

- a) zadeklarować macierz diagonalną stopnia 10, która na przekątnej ma same trójki,
- b) zadeklarować macierz diagonalną stopnia 13, która na przekątnej ma liczby od 1 do 13,
- c) sprawdzić, czy następująca macierz jest macierzą nieosobliwą:

$$\begin{bmatrix} 2 & 0 & 0 & 4 \\ 0 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ -1 & 0 & 1 & 0 \end{bmatrix}$$

7. Proszę wczytać do pakietu R dane o nazwie *milkgene* dostępne w bibliotece PBImisc (lub w pliku *milkgene.txt*), a następnie:
- a) sprawdzić wymiar i strukturę wczytanych danych
 - b) obliczyć, ile obserwacji zebrano dla każdej z laktacji (np. u ilu krów zmierzono wydajność z drugiej laktacji)
 - c) wyznaczyć wartość nowej zmiennej – zawartości procentowej tłuszczu w mleku
 - d) przyłączyć nową zmienną do istniejącej już ramki danych i nadać jej nazwę *fat_perc*
 - e) zapisać powstałą ramkę danych do pliku o nazwie *milkgene1.txt*
8. Proszę zapoznać się z działaniem funkcji *stat.desc*. W jakim pakiecie można ją znaleźć? Jakie jest jej zastosowanie? Do analizy których zmiennych z danych *milkgene* można użyć? Proszę zastosować funkcję do analizy tych zmiennych.
9. Proszę napisać funkcję o nazwie *RelMatrixA*, która tworzy macierz spokrewnień na podstawie informacji rodowodowej. Informacja rodowodowa ma mieć postać tabeli, której pierwszą kolumnę stanowią numery wszystkich osobników oraz dwie kolejne kolumny - odpowiednio numery ojca i matki tych osobników. Jeżeli rodzic lub rodzice osobnika są nieznani, należy nadać im numer zero. Tabela musi zawierać tyle wierszy, ile jest osobników. Funkcja powinna zapisywać utworzoną macierz do pliku *A.txt*. Proszę utworzyć tabelę danych rodowodowych na podstawie przykładu podanego na wykładzie oraz przetestować z jej wykorzystaniem funkcję *RelMatrixA*.
10. Proszę podać przykład próby pochodzącej z populacji, którą można przedstawić z zastosowaniem:
- a) klasyfikacji pojedynczej z ortogonalnym układem próby,
 - b) klasyfikacji krzyżowej dwukierunkowej bez interakcji,
 - c) klasyfikacji hierarchicznej dwustopniowej,
- a następnie opisać modele wskazanych klasyfikacji oraz przedstawić je w zapisie macierzowym.

11. Proszę zilustrować obliczeniami w pakiecie R przykład z wykładu dotyczący wykorzystania metody BLUP z modelem ojcowskim.

nr stada	nr ojca	nr córki	wydajność mleka
1	1	111	3677
1	1	112	4161
1	1	113	3505
1	1	114	3904
1	3	131	3957
1	3	132	3447
2	1	211	3534
2	2	221	2941
2	2	222	3366
2	2	223	3755
2	3	231	2318
2	3	232	2730
2	3	233	3629
2	3	234	3158

12. Proszę zilustrować obliczeniami w pakiecie R przykład z wykładu dotyczący wykorzystania metody BLUP z modelem osobniczym.

13. Proszę zastosować metodę BLUP z modelem osobniczym z wykorzystaniem następujących danych:

nr zwierzęcia	nr ojca	nr matki	y	rok
1	-	-	-	-
2	-	-	225	1
3	-	-	200	1
4	-	-	255	1
5	1	3	250	2
6	1	3	198	2
7	2	4	245	2
8	2	4	260	2
9	2	4	235	2