

Solución 1er Trabajo Encargado

AUTHOR

Genética General Avanzada

Para el presente trabajo use datos de pedigrí `pedigri_Merino` y de fenotipos `fenotipo_Merino.txt` que fueron simulados para ovinos.

Preguntas

1. ¿Cuántos animales contiene el archivo de pedigrí: *pedigri_Merino*? (2 puntos)

Limpiamos objetos y figuras

```
rm(list = ls())  
graphics.off()
```

Activamos librerías esenciales

```
library(tidyverse)  
library(kableExtra)
```

Indicamos el path de directorio de trabajo

```
setwd("~/XXX/XXX/Trabajo1")
```

```
pedigri <- read.table("pedigri_Merino.txt", header = T)  
nrow(pedigri)
```

```
[1] 33800
```

```
head(pedigri,3)
```

	IIId	FId	MIId
1	Merino17191	Merino15785	Merino15662
2	Merino17192	0	0
3	Merino17193	Merino15541	Merino15696

2. ¿Calcule es el promedio de `d_fibra`(Diametro de fibra) por cada año del archivo de fenotipo: `fenotipo_Merino.txt` ? (2 puntos)

```
fenotipo <- read.table("fenotipo_Merino.txt", header = T, na.strings = "0")  
head(fenotipo,3)
```

	IIId	sex	Anho	Population	d_fibra	p_vellongras	peso_vivo
1	Merino34991	M	2020	Ayaviri	11.31	2.46	42.37

2 Merino34992	F 2020	Ayaviri	12.46	NA	45.70
3 Merino34993	M 2020	Ayaviri	12.63	3.29	46.61

```
fenotipo %>%
  drop_na(d_fibra) %>%
  group_by(Anho) %>%
  summarise(Número = n(),
            Promedio = mean(d_fibra)) %>%
  rename(Año = Anho) %>%
  kbl()
```

	Año	Número	Promedio
	2020	2896	13.74734
	2021	2885	13.20395
	2022	2903	12.70364
	2023	2867	12.13486
	2024	2894	11.59051

3. ¿Calcule el promedio de `peso_vivo(Peso vivo del animal)` por cada año del archivo de fenotipo: `fenotipo_Merino.txt` ? (2 puntos)

```
fenotipo %>%
  drop_na(peso_vivo) %>%
  group_by(Anho) %>%
  summarise(Número = n(),
            Promedio = mean(peso_vivo)) %>%
  rename(Año = Anho) %>%
  kbl()
```

	Año	Número	Promedio
	2020	2537	44.59425
	2021	2544	44.61241
	2022	2537	44.61755
	2023	2599	44.62366
	2024	2512	44.75117

4. ¿Cual es el promedio general de los valores genéticos predichos (EBV) del diámetro de fibra ? (2 puntos)

Para correr el modelo use el siguiente modelo :

$$\text{caracteristica} = \text{sexo} + \text{anho} + \text{poblacion} + \text{animal} + e$$

Creamos archivo de parametros: `param_BLUP_df.txt`

```
DATAFILE
fenotipo_Merino.txt
SKIP_HEADER
1
TRAITS
5      # d_fibra
FIELDS_PASSED TO OUTPUT

WEIGHT(S)

RESIDUAL_VARIANCE
0.6
EFFECT
2 cross alpha # sex
EFFECT
3 cross alpha # Anho
EFFECT
4 cross alpha # Population
EFFECT
1 cross alpha # animal
RANDOM
animal
OPTIONAL

FILE
pedigri_Merino.txt
SKIP_HEADER
1
PED_DEPTH
0
(CO)VARIANCES
0.4
OPTION method BLUP
OPTION store_accuracy 4 orig
OPTION acctype 1.0
OPTION inbreeding_method 1
```

Corremos los programas BLUPF90

```
system("./renumf90 param_BLUP_df.txt")
system("./blupf90+ renf90.par")
```

Copiamos el archivo de salida `acc_bf90` con otro nombre para evitar confusiones. Si el comando no funciona en Windows, lo puede copiar de la manera tradicional

```
mv acc_bf90 acc_bf90_df
```

```
soluciones_df <- read.table("acc_bf90_df", header = F)
```

```
soluciones_df %>%
  select(-V1, -V2) %>%      # Elimina las dos primeras columnas
  rename(id = V3,           # Renombra V3 a "id"
         EBV_df = V4,       # Renombra V4 a "EBV_df"
         r2_df = V5) %>%    # Renombra V5 a "r2_df"
  mutate(DEP_df = EBV_df/2,  # Divide EBV_df entre dos y crea DEP_df
         r_df = sqrt(r2_df)) %>% # Raiz cuadrada de confiabilidad
  summarise(Promedio_DF = mean(EBV_df)) %>% # promedio de EBV_df
  kbl()
```

Promedio_DF

0.1128742

5. Hacer un ranking de los 10 machos con mayor DEP para diámetro de fibra (2 puntos)

Leemos el archivo de soluciones

```
soluciones_df <- read.table("acc_bf90_df", header = F)
```

Usamos un paquete de R para obtener la información del sexo de cada individuo

```
pedigri <- read.table("pedigri_Merino.txt", header = T, na.strings = "0")
library(optiSel)
total_pedigri <- prePed(pedigri)
info_sexo <- total_pedigri[, c(1,4)]
```

Combinamos el objeto de soluciones y el de información del sexo

```
sol_df_sex <- merge(soluciones_df, info_sexo, by.x = "V3", by.y = "Indiv")
```

```
sol_df_sex %>%
  filter(Sex == "male") %>% # Selecciona solo machos
  select(-V1, -V2) %>%      # Elimina las dos primeras columnas
  rename(id = V3,           # Renombra V3 a "id"
         EBV_df = V4,       # Renombra V4 a "EBV_df"
         r2_df = V5) %>%    # Renombra V5 a "r2_df"
  mutate(DEP_df = EBV_df / 2, # Divide EBV_df entre dos y crea DEP_df
         r_df = sqrt(r2_df),  # Saca raíz cuadrada de r2_df
         ) %>% # Crea columna de Ranking
  arrange(DEP_df) %>%
  mutate(Ranking = 1:n()) %>%
  head(10) %>%
  kbl()
```

id	EBV_df	r2_df	Sex	DEP_df	r_df	Ranking
Merino48589	-3.043859	0.8007	male	-1.521929	0.8948184	1

id	EBV_df	r2_df	Sex	DEP_df	r_df	Ranking
Merino48375	-3.006438	0.8113	male	-1.503219	0.9007219	2
Merino47937	-2.932362	0.8456	male	-1.466181	0.9195651	3
Merino48733	-2.579869	0.7912	male	-1.289935	0.8894942	4
Merino48681	-2.516906	0.7319	male	-1.258453	0.8555115	5
Merino45165	-2.468464	0.8257	male	-1.234232	0.9086804	6
Merino48221	-2.390137	0.7147	male	-1.195069	0.8453993	7
Merino47991	-2.372532	0.8295	male	-1.186266	0.9107689	8
Merino48391	-2.367686	0.7673	male	-1.183843	0.8759566	9
Merino49245	-2.342303	0.7555	male	-1.171151	0.8691950	10

6. Hacer un rankig de los 10 machos con mayor valor de precision (**accuracy**) para diametro de fibra (2 puntos)

```
sol_df_sex %>%
  filter(Sex == "male") %>% # Selecciona solo hembras
  select(-V1, -V2) %>%      # Elimina las dos primeras columnas
  rename(id = V3,           # Renombra V3 a "id"
         EBV_df = V4,       # Renombra V4 a "EBV_df"
         r2_df = V5) %>%    # Renombra V5 a "r2_df"
  mutate(DEP_df = EBV_df / 2, # Divide EBV_df entre dos y crea DEP_df
         r_df = sqrt(r2_df), # Saca raíz cuadrada de r2_df
         ) %>% # Crea columna de Ranking
  arrange(-r_df) %>%
  mutate(Ranking = 1:n()) %>%
  head(10) %>%
  kbl()
```

id	EBV_df	r2_df	Sex	DEP_df	r_df	Ranking
Merino44983	-1.7812115	0.8937	male	-0.8906058	0.9453571	1
Merino38467	-1.0960523	0.8890	male	-0.5480262	0.9428680	2
Merino44841	-1.7909625	0.8877	male	-0.8954812	0.9421783	3
Merino35353	-0.5057201	0.8873	male	-0.2528601	0.9419660	4
Merino38503	-1.0265658	0.8855	male	-0.5132829	0.9410101	5
Merino38569	-0.5029243	0.8830	male	-0.2514622	0.9396808	6
Merino35193	-0.5067751	0.8827	male	-0.2533876	0.9395212	7
Merino38949	-1.0738712	0.8815	male	-0.5369356	0.9388823	8
Merino41767	-2.2065847	0.8807	male	-1.1032924	0.9384562	9
Merino38449	-0.9974600	0.8791	male	-0.4987300	0.9376033	10

7. Hacer un ranking de las 10 hembras con mayor DEP para peso vellón (2 puntos)

Recordando que:

$$DEP = \frac{1}{2}EBV$$

Donde

- DEP = Diferencia esperada de la progenie, *expected progeny difference*
- EBV = Valor de cría, *estimated breeding value*

Creamos archivo de parametros: `param_BLUP_pve.txt`

```
DATAFILE
fenotipo_Merino.txt
SKIP_HEADER
1
TRAITS
6      # p_vellongras
FIELDS_PASSED TO OUTPUT

WEIGHT(S)

RESIDUAL_VARIANCE
0.6
EFFECT
2 cross alpha # sex
EFFECT
3 cross alpha # Anho
EFFECT
4 cross alpha # Population
EFFECT
1 cross alpha # animal
RANDOM
animal
OPTIONAL

FILE
pedigri_Merino.txt
SKIP_HEADER
1
PED_DEPTH
0
(CO)VARIANCES
0.4
OPTION method BLUP
OPTION store_accuracy 4 orig
OPTION acctype 1.0
OPTION inbreeding_method 1
```

Corremos los programas BLUPF90

```
system("./renumf90 param_BLUP_pve.txt")
system("./blupf90+ renf90.par")
```

```
mv acc_bf90 acc_bf90_pve
```

Leemos el archivo de soluciones

```
soluciones_pve <- read.table("acc_bf90_pve", header = F)
```

Usamos un paquete de R para obtener la informacion del sexo de cada individuo

```
pedigri <- read.table("pedigri_Merino.txt", header = T, na.strings = "0")
library(optiSel)
total_pedigri <- prePed(pedigri)
info_sexo <- total_pedigri[, c(1,4)]
```

Combinamos el objeto de soluciones y el de informacion del sexo

```
sol_pve_sex <- merge(soluciones_pve, info_sexo, by.x = "V3", by.y = "Indiv")
```

```
sol_pve_sex %>%
  filter(Sex == "female") %>% # Selecciona solo hembras
  select(-V1, -V2) %>%       # Elimina las dos primeras columnas
  rename(id = V3,            # Renombra V3 a "id"
         EBV_pve = V4,      # Renombra V4 a "EBV_df"
         r2_pve = V5) %>%   # Renombra V5 a "r2_df"
  mutate(DEP_pve = EBV_pve / 2, # Divide EBV_df entre dos y crea DEP_df
         r_pve = sqrt(r2_pve), # Saca raíz cuadrada de r2_df
         ) %>% # Crea columna de Ranking
  arrange(-DEP_pve) %>%
  mutate(Ranking = 1:n()) %>%
  head(10) %>%
  kbl()
```

id	EBV_pve	r2_pve	Sex	DEP_pve	r_pve	Ranking
Merino40694	2.280080	0.6078	female	1.1400399	0.7796153	1
Merino37418	2.238410	0.5750	female	1.1192048	0.7582875	2
Merino40866	2.182472	0.6278	female	1.0912359	0.7923383	3
Merino44048	2.153941	0.6059	female	1.0769703	0.7783958	4
Merino43402	2.088549	0.6210	female	1.0442746	0.7880355	5
Merino38316	2.019589	0.6089	female	1.0097945	0.7803204	6
Merino46104	1.980772	0.5970	female	0.9903861	0.7726578	7
Merino44454	1.916650	0.5859	female	0.9583251	0.7654410	8
Merino41676	1.900877	0.6333	female	0.9504383	0.7958015	9
Merino46970	1.880490	0.6163	female	0.9402452	0.7850478	10

8. ¿Cual es el macho con la mayor DEP para diametro de fibra y cómo se interpreta ese valor? (2 puntos)

Leemos el archivo de soluciones

```
soluciones_df <- read.table("acc_bf90_df", header = F)
```

Usamos un paquete de R para obtener la informacion del sexo de cada individuo

```
pedigri <- read.table("pedigri_Merino.txt", header = T, na.strings = "0")
library(optiSel)
total_pedigri <- prePed(pedigri)
info_sexo <- total_pedigri[, c(1,4)]
```

Combinamos el objeto de soluciones y el de información del sexo

```
sol_df_sex <- merge(soluciones_df, info_sexo, by.x = "V3", by.y = "Indiv")
```

```
sol_df_sex %>%
  filter(Sex == "male") %>%      # Selecciona solo hembras
  select(-V1, -V2) %>%          # Elimina las columnas V1 y V2
  rename(id = V3,                # Renombra V3 a "id"
         EBV_df = V4,           # Renombra V4 a "EBV_df"
         r2_df = V5) %>%        # Renombra V5 a "r2_df"
  mutate(DEP_df = EBV_df / 2,    # Divide EBV_df entre dos y crea DEP_df
         r_df = sqrt(r2_df),    # Saca raíz cuadrada de r2_df
         ) %>%
  arrange(DEP_df) %>%
  mutate(Ranking = 1:n()) %>%   # Crea columna de Ranking
  head(1) %>%
  kbl()
```

id	EBV_df	r2_df	Sex	DEP_df	r_df	Ranking
Merino48589	-3.043859	0.8007	male	-1.521929	0.8948184	1

El individuo **Merino48589** tiene la mejor DEP para diámetro de fibra con un valor de : **-1.521929 um**. Este valor indica que se espera que en media la progenie de ese individuo tendrá -1.5 um de diámetro de lana comparado con la media poblacional

9. ¿ Cuanto es la heredabilidad de diametro de fibra? (2 puntos)

Creamos al archivo de parametros **param_VCE_df.txt** con el siguiente detalle:

```
DATAFILE
fenotipo_Merino.txt
SKIP_HEADER
1
```



```

TRAITS
5      # diametro fibra
FIELDS_PASSED TO OUTPUT

WEIGHT(S)

RESIDUAL_VARIANCE
0.30
EFFECT
2 cross alpha # sex
EFFECT
3 cross alpha # anho
EFFECT
4 cross alpha # poblacion
EFFECT
1 cross alpha # animal
RANDOM
animal
OPTIONAL

FILE
pedigri_Merino.txt
SKIP_HEADER
1
PED_DEPTH
0
(CO)VARIANCES
0.70
OPTION method VCE
OPTION EM-REML 100
OPTION se_covar_function H2_t1 G_4_4_1_1/(G_4_4_1_1+R_1_1)

```

Corremos los programas BLUPf90

```

system("./renumf90 param_VCE_df.txt")
system("./blupf90+ renf90.par")

```

```

Final Estimates
Genetic variance(s) for effect  4
0.95423
Residual variance(s)
0.47207
inverse of AI matrix (Sampling Variance)
0.11182E-02 -0.50652E-03
-0.50652E-03  0.35712E-03
Correlations from inverse of AI matrix
1.0000      -0.80153
-0.80153      1.0000
SE for G
0.33440E-01
SE for R
0.18898E-01

Sampling variances of covariances function of random effects (n=10000)

```

```
H2_t1 - Function: g_4_4_1_1/(g_4_4_1_1+r_1_1)
Mean:    0.66903
Sample Mean:  0.66874
Sample SD:    0.15900E-01
elapsed time  1.0513306E-02
```

La h^2 de diámetro de fibra = 0.67 ± 0.016

10. ¿ Cuanto es la heredabilidad de diametro de fibra? (2 puntos)

Creamos al archivo de parametros `param_VCE_pve.txt` con el siguiente detalle:

```
DATAFILE
fenotipo_Merino.txt
SKIP_HEADER
1
TRAITS
6      # peso vellon
FIELDS_PASSED TO OUTPUT

WEIGHT(S)

RESIDUAL_VARIANCE
0.30
EFFECT
2 cross alpha # sex
EFFECT
3 cross alpha # anho
EFFECT
4 cross alpha # poblacion
EFFECT
1 cross alpha # animal
RANDOM
animal
OPTIONAL

FILE
pedigri_Merino.txt
SKIP_HEADER
1
PED_DEPTH
0
(CO)VARIANCES
0.70
OPTION method VCE
OPTION EM-REML 200
OPTION se_covar_function H2_t1 G_4_4_1_1/(G_4_4_1_1+R_1_1)
```

Corremos los programas BLUPf90

```
system("./renumf90 param_VCE_pve.txt")
system("./blupf90+ renf90.par")
```

Final Estimates

Genetic variance(s) for effect 4

0.51609

Residual variance(s)

1.3276

inverse of AI matrix (Sampling Variance)

0.14309E-02 -0.87941E-03

-0.87941E-03 0.99947E-03

Correlations from inverse of AI matrix

1.0000 -0.73535

-0.73535 1.0000

SE for G

0.37828E-01

SE for R

0.31614E-01

Sampling variances of covariances function of random effects (n=10000)

H2_t1 - Function: g_4_4_1_1/(g_4_4_1_1+r_1_1)

Mean: 0.27993

Sample Mean: 0.27966

Sample SD: 0.18756E-01

elapsed time 8.6975098E-03

La h^2 de peso vellón es = 0.28 ± 0.0187