problem-set-4

May 29, 2021

Ethan Holleman

1 Problem 1

```
[1]: import math, statistics
     import pandas as pd
```

1.1 Part A

What is Ghost's breeding value?

```
[2]: pop_mean = 900
     ghost_myw = 1200
     progeny_diff = 850 - pop_mean
     ghost_breeding_value = 2 * progeny_diff
     print('Ghost breeding value estimate:', ghost_breeding_value)

     # should be -100 lbs using 850
```

```
Ghost breeding value estimate: -100
```

1.2 Part B

One explanation for why Ghost could have had a high yearling weight but produce offspring with comparatively low mean yearling weight are temporary environmental effects which obscured the potentially positive effect Ghost would have had on yearling weight of his offspring. It is unlikely permanent environmental effects played as much of a factor as these effects could not be potentially attributed to effects specific to Ghost's offspring. Additionally, the yearling weight of Ghost's offspring may have been more greatly influence by inter and / or intra locus interactions; dominance or epistasis effects. These affects can be cumulatively summarized as the "allele combination value" and contribute to phenotype in a way that is independent of breeding value or environmental effects.

1.3 Part C

$ACV = 0$

$P = M + BV + ACV + E$

$1200 = 900 + 600 + 0 + E$

-300 = E

2 Problem 2

2.1 Part A

```
[3]: HH_len, Hh_len, hh_len = 100, 100, 150
     H_freq = 0.6
     h_freq = 1 - H_freq
     len_dict = {'HH': HH_len, 'Hh': Hh_len, 'hh': hh_len}
```

```
[4]: HH_freq = H_freq**2
     Hh_freq = 2 * H_freq * h_freq
     hh_freq = h_freq ** 2
     print(HH_freq, Hh_freq, hh_freq)
     assert sum((HH_freq, Hh_freq, hh_freq), 1)
```

```
0.36 0.48 0.16000000000000003
```

```
[5]: population_mean = sum((HH_len * HH_freq, Hh_freq * Hh_len, hh_freq * hh_len))
     print('Population mean: ', population_mean)
```

```
Population mean:  108.0
```

2.2 Part B

```
[6]: table_GV = [
         {'genotype': 'HH', 'Genotypic Value': HH_len - population_mean},
         {'genotype': 'Hh', 'Genotypic Value': Hh_len - population_mean},
         {'genotype': 'hh', 'Genotypic Value': hh_len - population_mean},
     ]
     table_GV = pd.DataFrame(table_GV)
     table_GV
```

```
[6]:    genotype  Genotypic Value
     0       HH             -8.0
     1       Hh             -8.0
     2       hh             42.0
```

2.3 Part C

```
[7]: HH_prog = {
         'HH': H_freq,
         'Hh': h_freq,
         'hh': 0
     }

     Hh_prog = {
```

```
        'HH': 0.5 * H_freq,
        'Hh': (0.5 * h_freq) + (0.5 * H_freq),
        'hh': 0.5 * h_freq
    }

    hh_prog = {
        'HH': 0,
        'Hh': H_freq,
        'hh': h_freq
    }
```

```
[8]: def prog_diff(pop_mean, prog, phenotype_dict):
         return sum([val * phenotype_dict[key] for key, val in prog.items()]) -␣
     ↪pop_mean


     HH_prog_diff = prog_diff(population_mean, HH_prog, len_dict)
     Hh_prog_diff = prog_diff(population_mean, Hh_prog, len_dict)
     hh_prog_diff = prog_diff(population_mean, hh_prog, len_dict)
```

```
[9]: table_BV = pd.DataFrame(
         [
             {'Genotype': 'HH', 'Breeding Value': HH_prog_diff*2},
             {'Genotype': 'Hh', 'Breeding Value': Hh_prog_diff*2},
             {'Genotype': 'Hh', 'Breeding Value': hh_prog_diff*2},
         ]
     )
     table_BV
```

```
[9]:   Genotype  Breeding Value
     0       HH           -16.0
     1       Hh             4.0
     2       Hh            24.0
```

2.4 Part D

If the frequency of the H allele increases to 0.8, then the frequency of h allele would decrease to 0.2. This would in turn effect the population mean which all other calculations are ultimately related to. Therefore *all* genotypic values can be expected to change.

3 Problem 3

PA stand for Producing Ability and can be thought of as the intrisitic capacity of an individual for production in the context of permeant eviornmental factors for repeated traits. This value considers the genotypic value of an individual and the permenant environmental effects acting on that indiviual. This metric is more useful when comparing individuals for production because an indiviudal with higher values will be less infulenced by temporary environmental effects. The

breeding value is not considered outside of its context as part of the genotypic value because when assessing PA we are not typically interested in the capacity of that individual to produce high quality offspring.

4 Problem 4

4.1 Part A

```
[10]: obs = (95, 110, 105, 95, 95)
      est_mean = statistics.mean(obs)
      print('Estimated mean:', est_mean)
```

Estimated mean: 100

4.2 Part B

```
[11]: est_var = statistics.variance(obs)
      print('Estimated varience:', est_var, 'mm^2')
```

Estimated varience: 50 mm^2

4.3 Part C

```
[12]: est_sd = statistics.stdev(obs)
      print('Estimated standard deviation:', round(est_sd, 2))
```

Estimated standard deviation: 7.07

5 Question 5

5.1 Part A

```
[13]: cov_bft = 20

      h2 = cov_bft / est_var   # covariance of breeding values / varience of␣
       ↪phenotypic values
      print('Estimated h^2:', h2)
```

Estimated h^2: 0.4

5.2 Part B

```
[14]: yb = 120
      bv_yb = (120 - est_mean) * h2
      print('Predicted breeding value of young boar:', bv_yb)
```

Predicted breeding value of young boar: 8.0

6 Question 6

6.1 Part A

```
[15]: ayw = 50   # lbs
      avh = 0.25   # average yearling heritibility
      br_br = 3.5   # best ram breeding value
      be_br = 4.5   # best ewe breeding value
      exp_br_champ = statistics.mean((br_br, be_br))
      print('Champ expected breeding value:', exp_br_champ)
```

Champ expected breeding value: 4.0

6.2 Part B

```
[16]: exp_ph_champ = 50 + exp_br_champ
      print('Champ expected year weight:', exp_ph_champ)
```

Champ expected year weight: 54.0

6.3 Part C

One explanation for Champ's lower than expected yearling weight could be that temporary environmental effects may be having a strong influence on phenotypes for yearling weight and therefore the breeding value of Champ's parents would have limited effect on Champ's actual phenotype.

Another possible explanation is that the reproducibility of yearling weight is low. Reproducibility is defined as the ratio of variance in production ability over phenotypic variance. PA itself is composed of breeding value, allele combination value and permanent environmental effects. Since we know the breeding value of Champ's parents to be high compared to the population it may be the case that allele combination value is dominating, minimizing the influence of breeding value.

6.4 Part D

```
[17]: champ_bv = (56 - 50) * avh
      print('Predicted Champ breeding value for yearling weight:', champ_bv)
```

Predicted Champ breeding value for yearling weight: 1.5

7 Question 7

7.1 Part A

```
[18]: t_my = 2000   # lbs
      t_fc = 0.1   # percent
      s_my = 1000   # lbs
      s_fc = 0.2   # percent

      # repeatability
```

```
r_my = 0.2
r_fc = 0.8

t_pa_milk, t_pa_fat = t_my * r_my, t_fc * r_fc
s_pa_milk, s_pa_fat = s_my * r_my, s_fc * r_fc

table = [
    {'Cow': 'Terri', 'PA Milk': t_pa_milk, 'PA Fat content': t_pa_fat},
     {'Cow': 'Sherry', 'PA Milk': s_pa_milk, 'PA Fat content': s_pa_fat}
]
pd.DataFrame(table)
```

[18]:
```
      Cow  PA Milk  PA Fat content
0   Terri    400.0            0.08
1  Sherry    200.0            0.16
```

7.2 Part B

I am more confident that Sherry is the better fat content producer because repeatibility is much higher for percent fat content, meaning that previous measures of milk fat content for a given cow should be more predictive of future milk fact content when compared to milk yield. Since we are only making one observation in time, we are more confident that our single observation is more representative of all possible observations when the reproducibility is higher.

7.3 Part C

$$\sigma_{pa}^2 = \sigma_{bv}^2 + \sigma_{acv}^2 + \sigma_{ep}^2 \tag{1}$$

$$\sigma_g^2 = \sigma_{bv}^2 + \sigma_{acv}^2 \tag{2}$$

$$\sigma_{ep}^2 = 0 \tag{3}$$

$$\sigma_{pa}^2 = \sigma_{bv}^2 + \sigma_{acv}^2 + 0 \tag{4}$$

$$\sigma_g^2 = \sigma_{pa}^2 \tag{5}$$

If milk production and fat content were not influenced by permanent environmental effects then broad sense heritability would equal reproducibility. Therefore broad sense heritability would be 0.2 and 0.8 for milk production and milk fat percentage respectively.