Simulation and parametric bootstrap

Data and goal

Data on 497 performances by 37 undergraduate music majors (between 2 and 15 performances were measured for each musician). Each row in the data represents one performance:

- `id`: a unique identifier for the musician
- `na`: negative affect score (a measure of anxiety)
- `large`: whether the musician was performing as part of a large ensemble (large = 1), or as part of a small ensemble or solo (large = 0)
- `audience`: who attended (Instructor, Public, Students, or Juried)

Research question: Is there a difference in anxiety between large and small ensemble performances, after accounting for audience type, and accounting for systematic variation between musicians?

Models

Full model:

$$Anxiety_{ij} = \beta_0 + \beta_1\ JuriedPerformance_{ij} + \beta_2\ PublicPerformance_{ij} + \beta_3\ StudentPerformance_{ij} + \beta_4\ LargeEnsemble_{ij} + u_i + \varepsilon_{ij}$$

$u_i \overset{iid}{\sim} N(0, \sigma_u^2), \varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2).$

> We want to test whether there is a difference between large and small ensemble performances. What is the reduced model?

Models

Full model:

$$Anxiety_{ij} = \beta_0 + \beta_1\ JuriedPerformance_{ij} + \beta_2\ PublicPerformance_{ij} + \beta_3\ StudentPerformance_{ij} + \beta_4\ LargeEnsemble_{ij} + u_i + \varepsilon_{ij}$$

$u_i \overset{iid}{\sim} N(0, \sigma_u^2), \varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2).$

Reduced model:

$$Anxiety_{ij} = \beta_0 + \beta_1\ JuriedPerformance_{ij} + \beta_2\ PublicPerformance_{ij} + \beta_3\ StudentPerformance_{ij} + u_i + \varepsilon_{ij}$$

$u_i \overset{iid}{\sim} N(0, \sigma_u^2), \varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$

Fitting the models

```
m1 <- lmer(na ~ audience + large + (1|id),
           data = music)
m0 <- lmer(na ~ audience + (1|id), data = music)
```

> What test statistic should I calculate to compare the models?

LRT

```
m1 <- lmer(na ~ audience + large + (1|id),
           data = music)
m0 <- lmer(na ~ audience + (1|id), data = music)
```

Likelihood ratio test statistic:

```
as.numeric(2*(summary(m1)$logLik -
                summary(m0)$logLik))
```

```
## [1] 12.459
```

Parametric bootstrapping

Observed test statistic: 12.46

> How would I use parametric bootstrapping to calculate a p-value for this test statistic?

Simulating from the reduced model

```
summary(m0)
```

```
...
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  id       (Intercept)  5.599   2.366
##  Residual             20.852   4.566
## Number of obs: 497, groups:  id, 37
##
## Fixed effects:
##                                Estimate Std. Error t value
## (Intercept)                     14.9288     0.5560  26.849
## audienceJuried Recital           3.8268     0.8183   4.677
## audiencePublic Performance       0.9454     0.5452   1.734
## audienceStudent(s)               2.9242     0.6246   4.682
...
```

Simulating from the reduced model

```
re_new <- rnorm(n = 37, mean = 0,
                sd = sqrt(5.60))
noise_new <- rnorm(n = 497, mean = 0,
                   sd = sqrt(20.85))
fitted_values <- predict(m0, re.form=NA)

re_data <- data.frame(id = unique(music$id),
                      re = re_new) %>%
  right_join(dplyr::select(music, id), by = "id")

new_data <- data.frame(id = music$id,
                       audience = music$audience,
                       large = music$large,
                       na = fitted_values +
                         re_data$re +
                         noise_new)
```

Calculate a test statistic with simulated data

> How do I calculate a test statistic using the simulated data?

Calculate a test statistic with simulated data

> How do I calculate a test statistic using the simulated data?

```
m1_sim <- lmer(na ~ audience + large + (1|id),
               data = new_data)
m0_sim <- lmer(na ~ audience + (1|id),
               data = new_data)
```

```
as.numeric(2*(summary(m1_sim)$logLik -
               summary(m0_sim)$logLik))
```

```
## [1] 1.128246
```

Repeat many times!