# Assignment 2

Your Name

2025-09-03

## 1.

Below are two snippets from Materials and Methods sections from peer-reviewed papers. Identify the treatment structure and design structure for both, or mention if this information is unclear or unavailable.

### 1a.

From "Plasma concentrations of substance P and cortisol in beef calves after castration or simulated castration" - Coetzee et al. (2008) [link to paper]

**Materials and Methods**

**Animals**—Ten Angus-crossbred calves were used in the study. Calves were 4 to 6 months old and weighed approximately 250 kg. Calves were acquired from a livestock commission company in Kansas in April 2006. Scrotal circumference (measured by use of a tape measure specifically manufactured for that purpose[a]) for each calf ranged from 22 to 30 cm prior to study commencement. On arrival at our facility, the calves received a single dose of a 7-way clostridial vaccine[b] (administered SC), a modified-live vaccine against viral respiratory disease[c] (administered IM), and a single injection of florfenicol[d] (40 mg/kg, SC). Calves also were provided amprolium[e] in the drinking water (10 mg/kg) for treatment and prevention of coccidiosis during the period of intensive housing in the study. The protocol for this study was approved by the Institutional Animal Care and Use Committee at Kansas State University (protocol No. 2472).

(...)

**Assignment to groups**—Calves were blocked in pairs on the basis of the scrotal circumference measured prior to study commencement. Calves were ranked by ascending scrotal circumference and assigned a computer-generated[j] random number. In each pair, the calf with the highest random number was assigned to the castration group, whereas the other calf was designated as an uncastrated control calf (n = 5 calves/group). Mean ± SEM scrotal circumference was 25.80 ± 1.36 cm for the castration group and 26.40 ± 0.86 cm for the control group.

**Castration and simulated castration**—The study commenced at 7 AM. Castration or simulated castration was performed at 3-minute intervals. All castrations were performed by a single experienced veterinarian (BVL) to minimize variation.

The scrotum of each calf was washed with dilute chlorhexidine disinfectant. Castration was performed via an open surgical technique without the provision of local anesthesia. This was consistent with standard industry practices used at many intensive production facilities in the United States. After the scrotum was washed with disinfectant, it was incised with a sharp Newberry castrating knife.[k] The testes and spermatic cords were exteriorized by blunt dissection, and the scrotal fascia was stripped from each testis. A Henderson castration tool[k] was clamped across the entire spermatic cord immediately proximal to a testis. The tool was attached to a 6.0-V cordless variable-speed hand drill[l] with a 3/8-inch chuck. The drill was used to rotate the clamped spermatic cord at a slow to moderate speed in a clockwise direction; after an initial 5 or 6 revolutions, the drill speed was increased in accordance with the manufacturer's instructions until the torsion resulted in removal of the testis (approx 10 revolutions). The tightly torsed sealed segment of the cord then retracted into the abdomen. The same procedure was used to remove the second testis.

For each calf in the control group, the effect of manipulation associated with castration was simulated. The testes within the scrotum were firmly grasped, and ventral traction was applied for approximately 20 seconds.

Figure 1: From Coetzee et al. (2008)

**Treatment structure:** One-way treatment structure: castration & control.

**Design structure:** RCBD – blocks were pairs of calves with similar scrotal circumference.

**1b.**

From "Maize Kernel Weight Response to Postflowering Source–Sink Ratio" - Borras and Otegui (2001) [link to paper]

## MATERIALS AND METHODS

Two F1 commercial hybrids, DK752 and DK664 (Dekalb-Monsanto, Argentina), were sown in October during the growing seasons of 1998 to 1999 (Year 1) and 1999 to 2000 (Year 2) at Salto, Argentina (34°33′S, 60°33′W). The DK752 is a small-kernel hybrid (<250 mg kernel$^{-1}$) at 70 000 plants ha$^{-1}$, while the DK664 has large kernels (>250 mg kernel$^{-1}$) under the same growing conditions. Two plant populations were used, 3 plants m$^{-2}$ (low density) and 9 plants m$^{-2}$ (high density). They were selected in order to have contrasting vegetative (source) and grain (sink) biomass (Otegui, 1997). Treatments were arranged in a split-plot design with three replicates, where plant populations were the main plots and hybrids the subplots. Experimental plots were kept free of weeds and pests, and experienced no water or nutrient stress.

Three pollination treatments were performed in order to change the number of reproductive sinks per plant: restricted, hand, and open pollination. At least 10 plants per each hybrid stand density × pollination treatment replicate combination were tagged at random 15 d before silking, and were individually identified. The date of silking (first silks visible) of the apical and subapical ears was registered for each tagged plant. The restricted pollination treatment consisted of the controlled pollination of silks from apical ears. It was performed by bagging apical ears 2 d after they silked in order to decrease the number of pollinated ovaries. This manipulation avoided the negative effects observed when ears were cut in order to reduce KNP (Kiniry et al., 1990). In this treatment, the subapical ear was bagged prior to its silking to prevent its pollination. The hand-pollination treatment was aimed to synchronize the pollination of all exposed silks of the two ears 4 d after silking (DAS) of the apical ear (Frey, 1981; Cárcova et al., 2000). In this treatment, both ears of each plant were bagged before silking and were kept covered until 4 DAS, when fresh pollen of the same hybrid was added manually to all exposed silks. This treatment was done to improve KNP (Cárcova et al., 2000). The open-pollinated plants were never bagged and were used as control plants, with an expected intermediate KNP relative to the other two treatments. Plants from all treatments with irregular kernel set along the ear were discarded to avoid the confounded effect of unusually large kernels due to no space restriction.

Figure 2: From Borras and Otegui (2001)

**Treatment structure:** Three-way treatment structure: hybrid (DK752 and DK664), plant population (3 and 9 plants m$^2$), and pollination (restricted, hand, and open pollination).

**Design structure:** split-split-plot: plant pop in the whole plot, hybrid in the split-plot, pollination in the sub-subplot.

## 2.

The data in the code below correspond to a trial studying the effects of agronomic management and the timing of said agronomic management on plant height in rice. Studying plant height in rice is important because this trait is associated to lodging (i.e., when the plants are too tall and heavy and "fall"), which may reduce yields due to difficulty in harvesting the grain.

The scientists were able to divide the field into equally-sized blocks that fitted all treatments (i.e., *complete blocks*). Then, the scientists randomly allocated the management treatments to smaller areas within the blocks. After that, they assigned the timings to even smaller areas of the management treatments. From each block-management-timing treatment combination, they were able to measure plant height twice.

### 2a.

**What is the treatment structure and the design structure in this experiment?**

Answer: The treatment structure is a $7 \times 3$ factorial, and the design structure is a split-plot design with sub-sampling.

### 2b.

**Write the statistical model that best represents the data generating process.**

$$y_{ijkl}|b_k, w_{i(k)}, s_{j(ik)} \sim N(\mu_{ijkl}, \sigma^2_\varepsilon), \mu_{ijkl} = \eta_{ijkl} = \eta_0 + M_i + T_j + (MT)_{ij} + b_k + w_{i(k)} + s_{j(ik)},$$

where:

- $y_{ijkl}$ is the observation of the height of the $i$th management treatment, $j$th timing, $k$th block, and $l$th subsample out of that plot,
- $\mu_{ijkl}$ is the expected value,
- $\sigma^2_\varepsilon$ is the variance,
- $\eta_{ijkl}$ is the linear predictor,
- $\eta_0$ is the overall mean of the linear predictor,
- $M_i$ is the effect of the $i$th management treatment,
- $T_j$ is the effect of the $j$th timing treatment,
- $(MT)_{ij}$ is the interaction between the $i$th management treatment and the $j$th timing treatment,
- $b_k$ is the effect of the $k$th block, $b_k \sim N(0, \sigma^2_b)$
- $w_{i(k)}$ is the effect of the $i(k)$th whole plot, $w_{i(k)} \sim N(0, \sigma^2_w)$
- $s_{j(ik)}$ is the effect of the $j(ik)$th split plot, $w_{j(ik)} \sim N(0, \sigma^2_s)$.

### 2c.

Fit that model to the data, check model assumptions, report all variance components, and the estimated marginal means (including some measure of uncertainty) for all time-management combinations (i.e., 32 total combinations).

```
library(agridat)
data("gomez.splitplot.subsample")
df <- gomez.splitplot.subsample
```

```
library(tidyverse)
```

```
## Warning: package 'ggplot2' was built under R version 4.4.3
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.2     v tibble    3.2.1
```

```
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ---------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(lme4)
```

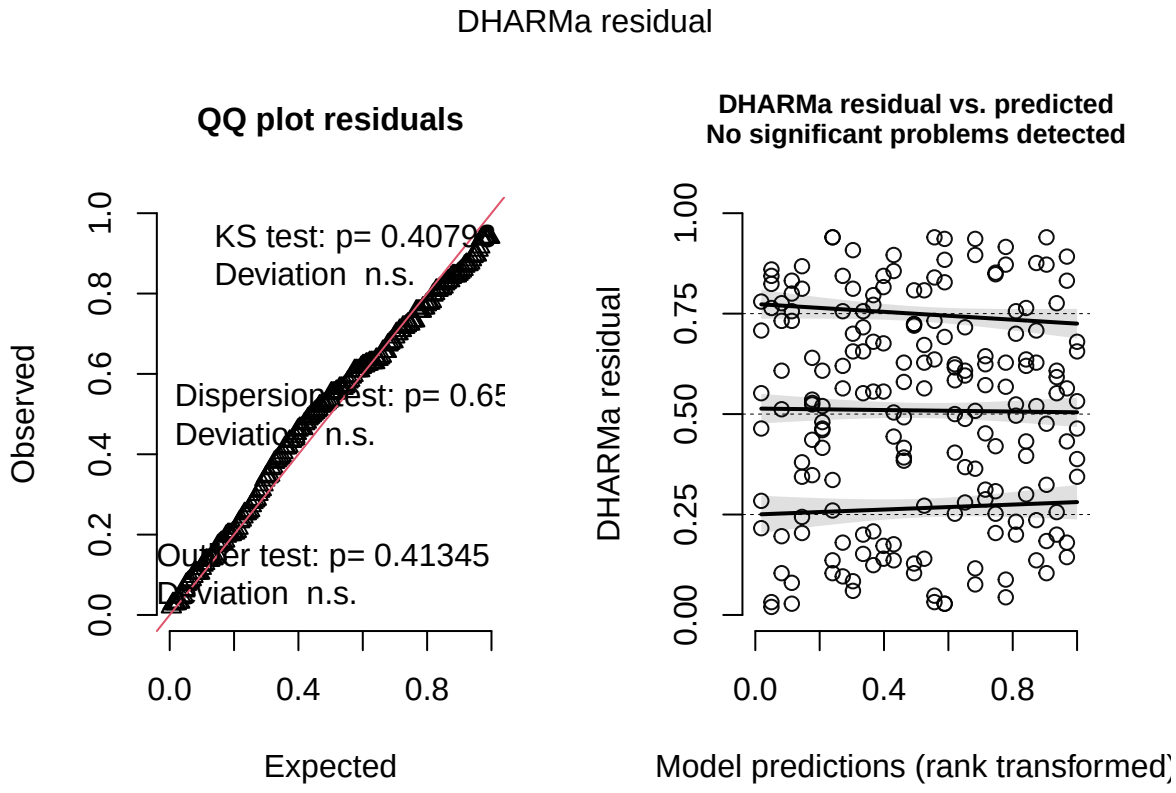
```
## Loading required package: Matrix
##
## Attaching package: 'Matrix'
##
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
```

```
library(emmeans)
```

```
## Welcome to emmeans.
## Caution: You lose important information if you filter this package's results.
## See '? untidy'
```

```
m <- lmer(height ~ time*manage + (1|rep/manage/time), data = df)
```

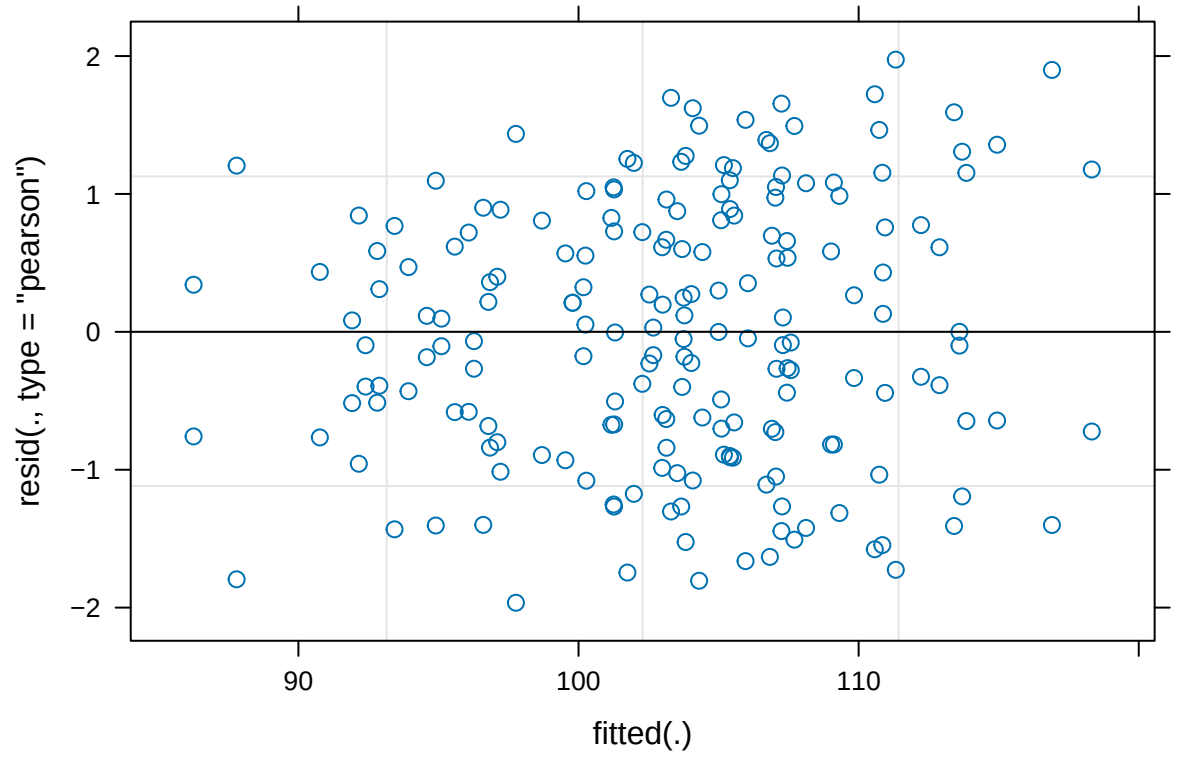
```
DHARMa::simulateResiduals(m, plot = T)
```




```
## Object of Class DHARMa with simulated residuals based on 250 simulations with refit = FALSE . See ?DH
##
```

```
## Scaled residual values: 0.3 0.048 0.176 0.172 0.58 0.28 0.028 0.628 0.388 0.252 0.028 0.136 0.116 0.
```

```
plot(m)
```




```
VarCorr(m)
```

```
##  Groups          Name        Std.Dev.
##  time:manage:rep (Intercept) 4.0433
##  manage:rep      (Intercept) 2.7422
##  rep             (Intercept) 4.3692
##  Residual                    1.3205
```

```
emmeans(m, ~time:manage)
```

```
##  time manage emmean   SE   df lower.CL upper.CL
##  T1   M1       108.0 3.82 8.63     99.3      117
##  T2   M1       111.2 3.82 8.63    102.5      120
##  T3   M1       105.8 3.82 8.63     97.1      114
##  T4   M1       109.3 3.82 8.63    100.6      118
##  T1   M2       103.6 3.82 8.63     94.9      112
##  T2   M2       105.3 3.82 8.63     96.6      114
##  T3   M2       101.8 3.82 8.63     93.1      111
##  T4   M2       100.2 3.82 8.63     91.5      109
##  T1   M3       102.7 3.82 8.63     94.0      111
##  T2   M3       104.0 3.82 8.63     95.2      113
##  T3   M3        98.1 3.82 8.63     89.4      107
##  T4   M3       103.2 3.82 8.63     94.4      112
##  T1   M4       101.8 3.82 8.63     93.1      111
```

```
##  T2   M4      108.1 3.82 8.63     99.4      117
##  T3   M4      100.5 3.82 8.63     91.8      109
##  T4   M4      111.0 3.82 8.63    102.3      120
##  T1   M5      103.0 3.82 8.63     94.3      112
##  T2   M5      105.2 3.82 8.63     96.5      114
##  T3   M5       99.9 3.82 8.63     91.2      109
##  T4   M5      101.8 3.82 8.63     93.1      111
##  T1   M6      104.6 3.82 8.63     95.9      113
##  T2   M6      108.3 3.82 8.63     99.6      117
##  T3   M6       98.4 3.82 8.63     89.7      107
##  T4   M6      107.0 3.82 8.63     98.3      116
##  T1   M7       98.3 3.82 8.63     89.6      107
##  T2   M7       99.5 3.82 8.63     90.8      108
##  T3   M7       97.1 3.82 8.63     88.4      106
##  T4   M7       98.2 3.82 8.63     89.5      107
##  T1   M8      103.4 3.82 8.63     94.7      112
##  T2   M8      105.2 3.82 8.63     96.5      114
##  T3   M8      100.3 3.82 8.63     91.6      109
##  T4   M8      104.5 3.82 8.63     95.7      113
##
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
```