

# Mixed-effect models simulations report

Mladen Jovanović

11/14/2020

This is summary report of the mixed-effects model simulation and estimation of the measurement random error.

Here we will create a function for simulation and plotting, which we will use later to plot various scenarios. The only **DIFFERENCE** with the previous simulation is that I am using first and last trial to LM, OLP, and Diff methods.

```
require(tidyverse)
require(bmbstats)
require(lme4)
require(cowplot)

mm_simulation <- function(n_athletes = 20,
                          n_trials = 5,
                          n_sim = 100,
                          intercept_min = 80,
                          intercept_max = 120,
                          slope_min = -5,
                          slope_max = 8,
                          random_min = 5,
                          random_max = 15,
                          cow_font = 6) {

  df <- expand_grid(
    simulation = seq(1, n_sim),
    athlete = LETTERS[seq(1, n_athletes)]
  ) %>%
  group_by(simulation) %>%
  mutate(
    intercept = runif(n(), intercept_min, intercept_max),
    slope = runif(n(), slope_min, slope_max),
    random = runif(n(), random_min, random_max)
  ) %>%
  ungroup() %>%
  expand_grid(
    trial = seq(0, n_trials - 1)
  ) %>%
  mutate(
    true_value = intercept + slope * trial,
    random_error = rnorm(n(), 0, random),
    observed_value = true_value + random_error
  )
}
```

```

# Plot
one_sim_data <- filter(df, simulation == 1)

gg1 <- ggplot(
  one_sim_data,
  aes(x = trial, y = true_value, group = athlete)
) +
  theme_cowplot(cow_font) +
  geom_line(alpha = 0.5) +
  geom_point(aes(y = observed_value), alpha = 0.5) +
  ylab(NULL)

gg2 <- gg1 + facet_wrap(~athlete)

# Analysis of the estimated random error
get_TE <- function(data) {
  data_wide <- data %>%
    pivot_longer(cols = -(1:6)) %>%
    pivot_wider(
      values_from = "value",
      names_from = c("name", "trial")
    )

  TE_diff <- sd(data_wide[[ncol(data_wide)]] - data_wide$observed_value_0) / sqrt(2)

  TE_lm <- summary(
    lm(data_wide[[ncol(data_wide)]] ~ data_wide$observed_value_0)
  )$sigma / sqrt(2)

  TE_olp <- OLP_regression(
    predictor = data_wide$observed_value_0,
    outcome = data_wide[[ncol(data_wide)]]
  )$rse / sqrt(2)

  mm0 <- lmList(
    observed_value ~ trial | athlete,
    data
  )

  mm1 <- lmer(
    observed_value ~ trial + (1 | athlete),
    data
  )

  mm2 <- lmer(
    observed_value ~ trial + (1 + trial | athlete),
    data
  )

  data$trial_nom <- factor(data$trial)

  mm3 <- lmer(

```

```

    observed_value ~ trial_nom + (1 | athlete),
    data
  )

  df <- tribble(
    ~model, ~TE,
    "Diff", TE_diff,
    "LM", TE_lm,
    "OLP", TE_olp,
    "Rand individual", summary(mm0)$RSE,
    "Rand int", summary(mm1)$sigma,
    "Rand int+slope", summary(mm2)$sigma,
    "Rand int (nom)", summary(mm3)$sigma
  )

  return(df)
}

# =====
TE <- df %>%
  group_by(simulation) %>%
  do(get_TE(.))

gg3 <- ggplot(TE, aes(y = TE, x = model)) +
  theme_cowplot(cow_font) +
  geom_boxplot(fill = "light grey") +
  coord_flip() +
  xlab(NULL)

# Add RE line plot
gg3 <- gg3 + geom_hline(yintercept = c(random_min, random_max), linetype = "dashed", color = "red")

plot_grid(gg2, gg3, nrow = 1)
}

```

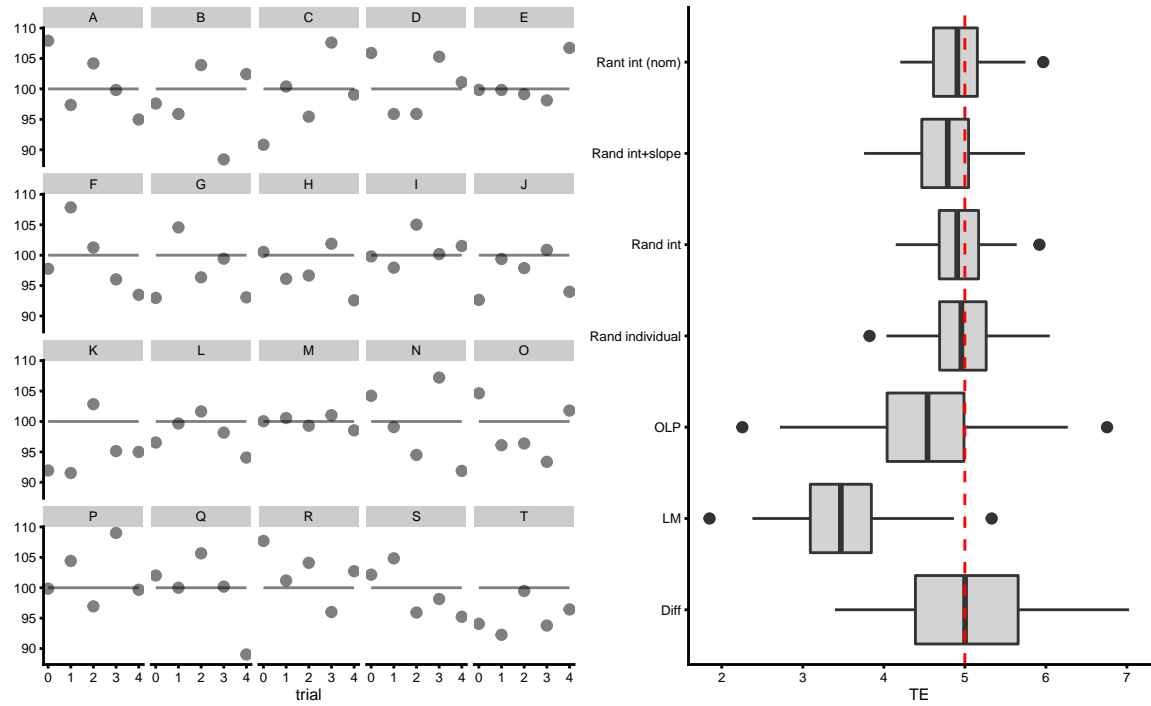
## Simple scenario

In this scenario, all participants have true values equal to 100 a.u. across 5 trials, with random error equal to 5 (0 a.u.)

```

mm_simulation(
  n_athletes = 20,
  n_trials = 5,
  n_sim = 100,
  intercept_min = 100,
  intercept_max = 100,
  slope_min = 0,
  slope_max = 0,
  random_min = 5,
  random_max = 5)

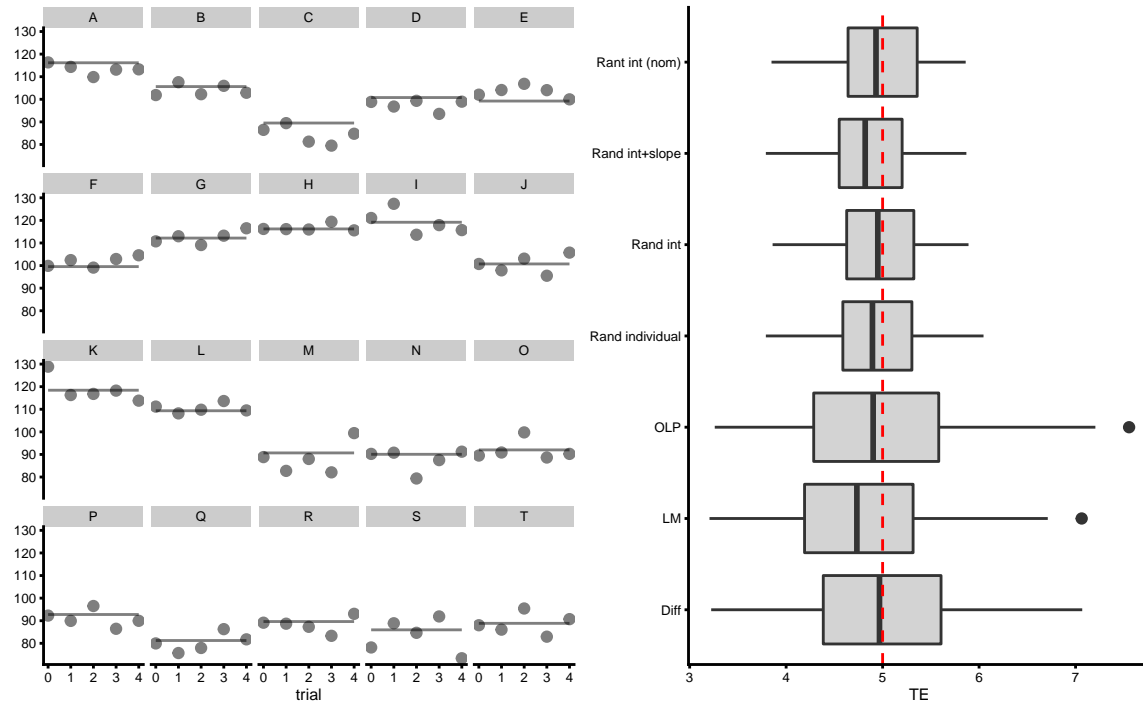
```



## Different intercepts

In this scenario, all participants have true values stable across 5 trials, with random error equal to 5 (0 a.u.)

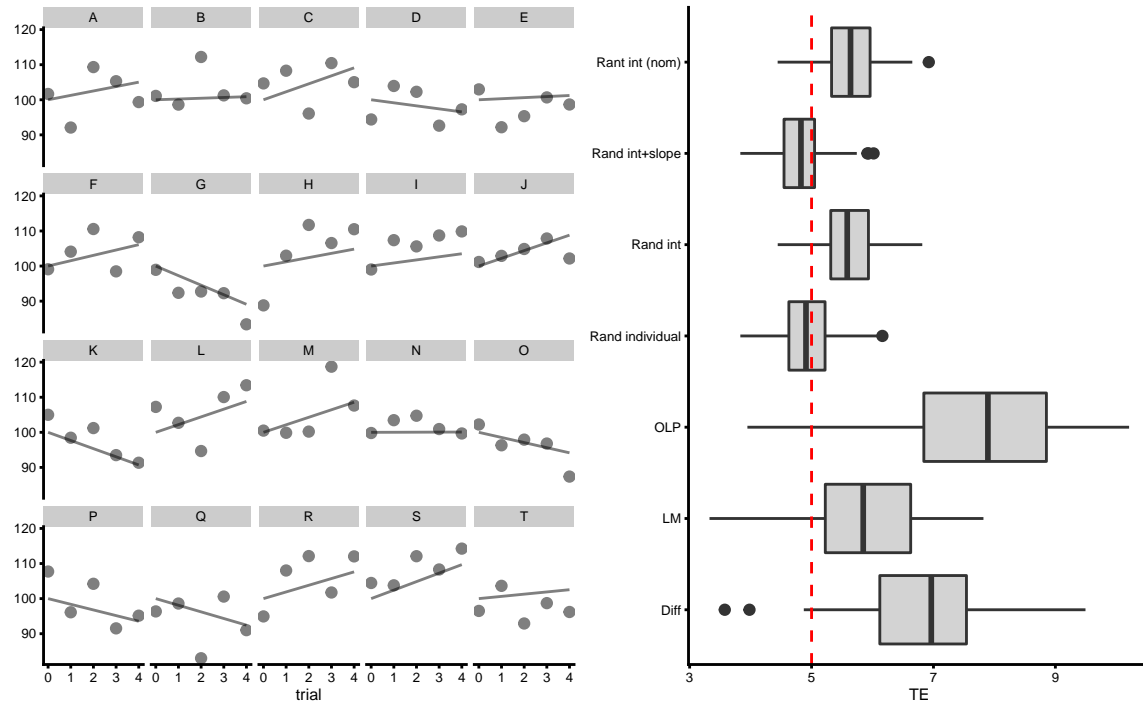
```
mm_simulation(
  n_athletes = 20,
  n_trials = 5,
  n_sim = 100,
  intercept_min = 80,
  intercept_max = 120,
  slope_min = 0,
  slope_max = 0,
  random_min = 5,
  random_max = 5)
```



## Different slopes

In this scenario, all participants have true values starting at 100 a.u., but different slopes across 5 trials, with random error equal to 5 (0 a.u.)

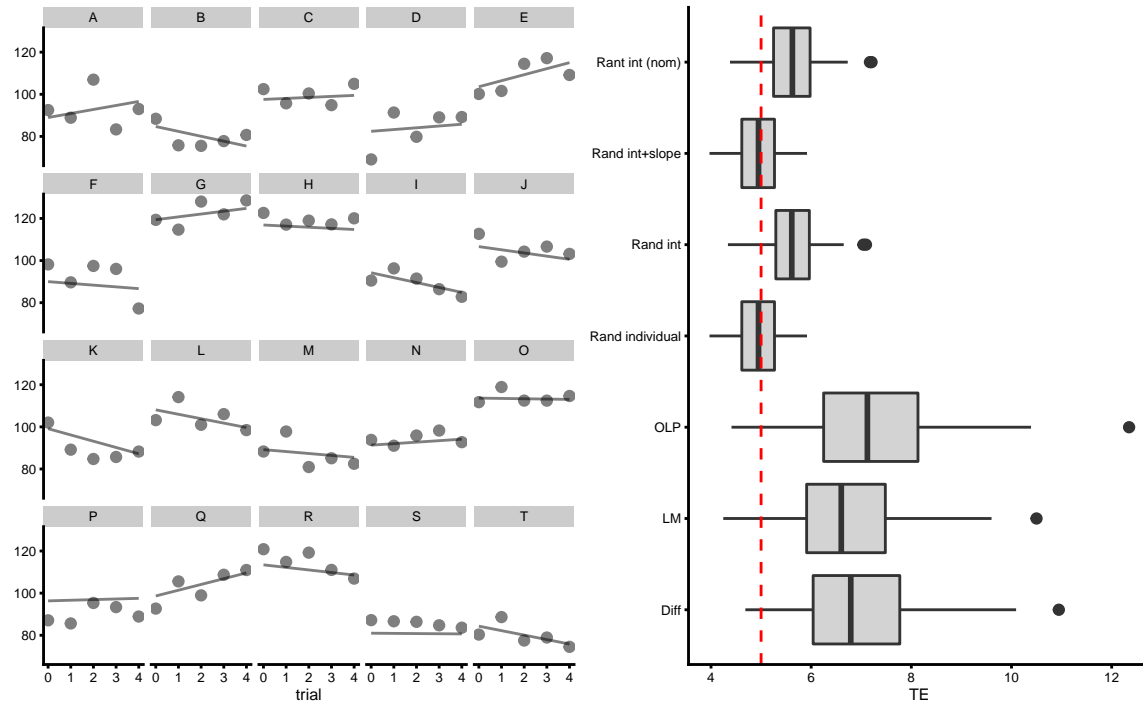
```
mm_simulation(
  n_athletes = 20,
  n_trials = 5,
  n_sim = 100,
  intercept_min = 100,
  intercept_max = 100,
  slope_min = -3,
  slope_max = 3,
  random_min = 5,
  random_max = 5)
```



## Different intercepts and different slopes

In this scenario, all participants have true values starting at different values, with different slopes across 5 trials, with random error equal to 5 (0 a.u.)

```
mm_simulation(
  n_athletes = 20,
  n_trials = 5,
  n_sim = 100,
  intercept_min = 80,
  intercept_max = 120,
  slope_min = -3,
  slope_max = 3,
  random_min = 5,
  random_max = 5)
```



## Different random errors

In this scenario, random errors vary across individuals on top of variable intercept and slope.

```
mm_simulation(
  n_athletes = 20,
  n_trials = 5,
  n_sim = 100,
  intercept_min = 80,
  intercept_max = 120,
  slope_min = -3,
  slope_max = 3,
  random_min = 1,
  random_max = 10)
```

