

Data Critique Simulations

2023-01-17

```
set.seed(1)
total_people_tested = 290
number_simulations = 1000

regression_sim = function(mean, variance_upper_bound, additional) {
  variances = runif(total_people_tested, 0.0001, variance_upper_bound)
  response = rnorm(total_people_tested, mean, variances)
  condition = factor(c(rep("solo",95), rep("peer",100), rep("adult",95)))
  df = data.frame(response = response, condition = condition)

  lm1 = lm(response ~ condition, data = df)
  peer_adult_p_value = t(summary(lm1)$coefficients[,4])[2]
  solo_adult_p_value = t(summary(lm1)$coefficients[,4])[3]

  # Now rerun the regression with peer as the reference level
  condition = relevel(condition, ref = "peer")
  reordered_df = data.frame(response = response, condition = condition)
  reordered_lm = lm(response ~ condition, data = reordered_df)
  solo_peer_p_value = t(summary(reordered_lm)$coefficients[,4])[3]

  # Only run this part of the code if you want to run the additional indifference
# point analysis based on the main regression and return all 6 p values
  if (additional) {
    actual = rnorm(total_people_tested, 1000, 200)
    D = rnorm(total_people_tested, 292, 20)
    df$indifference_point = actual / (1 + exp(df$response) * D)

    lm1 = lm(indifference_point ~ condition, data = df)
    pa2 = t(summary(lm1)$coefficients[,4])[2]
    sa2 = t(summary(lm1)$coefficients[,4])[3]

    condition = relevel(condition, ref = "peer")
    reordered_df = data.frame(indifference_point = df$indifference_point, condition = condition)
    reordered_lm = lm(indifference_point ~ condition, data = reordered_df)
    sp2 = t(summary(reordered_lm)$coefficients[,4])[3]

    return(c(peer_adult_p_value, solo_adult_p_value, solo_peer_p_value, pa2, sa2, sp2))
  }

  return(c(peer_adult_p_value, solo_adult_p_value, solo_peer_p_value))
}

extreme_results = 0
any_results = 0
```

```

simes_any_result = 0

for (i in 1:number_simulations) {

  comparison_values <- regression_sim(0.25, 0.1,FALSE)

  if (sum(comparison_values < 0.05) == 3) {
    extreme_results = extreme_results + 1
  }

  if (sum(comparison_values < 0.05) != 0) {
    any_results = any_results + 1
  }

}

extreme_results/number_simulations

```

```
## [1] 0
```

```
any_results/number_simulations
```

```
## [1] 0.125
```

```

extreme_results = 0
any_results = 0

for (i in 1:number_simulations) {

  risk_ps = regression_sim(0.25, 0.1,FALSE)
  log_k_indifference_ps = regression_sim(-6, 1,TRUE)
  all_ps = c(risk_ps, log_k_indifference_ps)

  if (sum(all_ps < 0.05) >= 5) {
    extreme_results = extreme_results + 1
  }

  if (sum(all_ps < 0.05) != 0) {
    any_results = any_results + 1
  }

}

extreme_results/number_simulations

```

```
## [1] 0.001
```

```
any_results/number_simulations
```

```
## [1] 0.29
```