Biostat 200C Midterm Makeup

Due June 3 @ 11:59PM

```
# if (!require("BiocManager", quietly = TRUE))
#    install.packages("BiocManager")
# BiocManager::install("msd16s")

# BiocManager::install("metagenomeSeq")
suppressMessages(library(metagenomeSeq))
library(msd16s)
library(tidyverse)
```

```
── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
✔ dplyr     1.1.4     ✔ readr     2.1.5
✔ forcats   1.0.0     ✔ stringr   1.5.1
✔ ggplot2   3.5.1     ✔ tibble    3.2.1
✔ lubridate 1.9.3     ✔ tidyr     1.3.1
✔ purrr     1.0.2
── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
✖ dplyr::combine()    masks Biobase::combine(), BiocGenerics::combine()
✖ tidyr::expand()     masks Matrix::expand()
✖ dplyr::filter()     masks stats::filter()
✖ dplyr::lag()        masks stats::lag()
✖ tidyr::pack()       masks Matrix::pack()
✖ ggplot2::Position() masks BiocGenerics::Position(), base::Position()
✖ tidyr::unpack()     masks Matrix::unpack()
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to
become errors
```

```
library(broom)
```

```
Warning: package 'broom' was built under R version 4.3.3
```

```
library(pscl)
```

```
Classes and Methods for R originally developed in the
Political Science Computational Laboratory
Department of Political Science
Stanford University (2002-2015),
by and under the direction of Simon Jackman.
hurdle and zeroinfl functions by Achim Zeileis.
```

```
library(VennDiagram)
```

```
Loading required package: grid
Loading required package: futile.logger
```

```
data(msd16s)
msd16s
```

```
MRexperiment (storageMode: environment)
assayData: 26044 features, 992 samples
  element names: counts
protocolData: none
phenoData
  sampleNames: 100259 100262 ... 602385 (992 total)
  varLabels: Type Country ... Dysentery (5 total)
  varMetadata: labelDescription
featureData
  featureNames: 54 94 ... 276421 (26044 total)
  fvarLabels: superkingdom phylum ... clusterCenter (10 total)
  fvarMetadata: labelDescription
experimentData: use 'experimentData(object)'
Annotation:
```

```
phenoData(msd16s)
```

```
An object of class 'AnnotatedDataFrame'
  sampleNames: 100259 100262 ... 602385 (992 total)
  varLabels: Type Country ... Dysentery (5 total)
  varMetadata: labelDescription
```

```
pheno_tbl <- rownames_to_column(pData(msd16s), var = "ID") %>% as_tibble()
```

```
p_case_control = pheno_tbl %>%
  count(Type) %>%
  mutate(prop = n/sum(n))

p_case <- with(p_case_control, prop[Type == "Case"])
p_control <- with(p_case_control, prop[Type == "Control"])
```

```
featureData(msd16s)
```

```
An object of class 'AnnotatedDataFrame'
  featureNames: 54 94 ... 276421 (26044 total)
  varLabels: superkingdom phylum ... clusterCenter (10 total)
  varMetadata: labelDescription
```

```
features <- fData(msd16s)
```

```
counts <- MRcounts(msd16s, norm = TRUE)
dim(counts)
```

```
[1] 26044   992
```

```
otu_id <- rownames(counts)
counts_tbl <- bind_cols(otu_id = otu_id, counts %>% as_tibble())
```

```
control_to_select <- pheno_tbl %>%
  filter(Type == "Control") %>%
  dplyr::select(ID) %>%
  pull()

con_sum <- counts_tbl %>%
  dplyr::select(all_of(control_to_select)) %>%
  rowSums()

case_to_select <- pheno_tbl %>%
  filter(Type == "Case") %>%
  dplyr::select(ID) %>%
  pull()

case_sum <- counts_tbl %>%
  dplyr::select(all_of(case_to_select)) %>%
  rowSums()

counts_tbl_filt <- bind_cols(counts_tbl,
                             con_sum = con_sum/length(control_to_select),
                             case_sum = case_sum/length(case_to_select)) %>%
  filter(con_sum >= 12 | case_sum >= 12)
```

```
otu_prevalence <- counts_tbl_filt %>%
  dplyr:: select(-otu_id) %>%
  mutate(across(everything(), ~ as.integer(. > 0))) %>%
  rowSums()

counts_tbl <- bind_cols(counts_tbl_filt, otu_prevalence = otu_prevalence) %>%
  filter(otu_prevalence >= 10) %>%
  dplyr:: select(-con_sum, -case_sum, -otu_prevalence)
```

```
counts_tbl_t <- counts_tbl %>%
  pivot_longer(cols= -1) %>%
  pivot_wider(names_from = "otu_id",values_from = "value") %>%
  rename(ID = name) %>%
  left_join(pheno_tbl, by = "ID")
```

Filter Data at Genus level

```
genus_counts = aggTax(msd16s, lvl = "genus", out = "matrix", norm = F)
genus_id <- rownames(genus_counts)
genus_counts_tbl <- bind_cols(genus_id = genus_id, genus_counts %>% as_tibble())
genus_counts_tbl
```

```
# A tibble: 164 × 993
   genus_id       `100259` `100262` `100267` `100274` `100275` `100277` `100291`
   <chr>             <dbl>    <dbl>    <dbl>    <dbl>    <dbl>    <dbl>    <dbl>
 1 Abiotrophia           0        0        0        0        0        0        1
 2 Acetobacter           0        0        0        0        0        0        0
 3 Acetonema             0        0        0        0        0        0        0
 4 Acidaminococc…        6        0        0       26        0       31        0
 5 Acidithiobaci…        0        0       79        0        0        0        0
 6 Acidovorax            0        0        0        0        0        0        0
 7 Acinetobacter         0        0        0        0        0        1        0
 8 Actinobacillus        0        0        0        0        0        0        0
 9 Actinomyces          11        0        1        0        0        0        0
10 Aeromonas             0        0        0        0        0        0        0
# ℹ 154 more rows
# ℹ 985 more variables: `100292` <dbl>, `100293` <dbl>, `100294` <dbl>,
#   `100298` <dbl>, `100303` <dbl>, `100317` <dbl>, `100320` <dbl>,
#   `100322` <dbl>, `100341` <dbl>, `100353` <dbl>, `100356` <dbl>,
#   `100361` <dbl>, `100365` <dbl>, `100395` <dbl>, `100401` <dbl>,
#   `100403` <dbl>, `100437` <dbl>, `100457` <dbl>, `100462` <dbl>,
#   `100470` <dbl>, `100476` <dbl>, `100489` <dbl>, `100491` <dbl>, …
```

```r
con_mean <- genus_counts_tbl %>%
  dplyr:: select(all_of(control_to_select)) %>%
  rowMeans()

case_mean <- genus_counts_tbl %>%
  dplyr:: select(all_of(case_to_select)) %>%
  rowMeans()

genus_tbl_filt <- bind_cols(genus_counts_tbl,
                            con_mean = con_mean,
                            case_mean = case_mean) %>%
  filter(con_mean >= 12 | case_mean >= 12)
```

```r
otu_prevalence <- genus_tbl_filt %>%
  dplyr:: select(-genus_id) %>%
  mutate(across(everything(), ~ as.integer(. > 0))) %>%
  rowSums()

genus_tbl <- bind_cols(genus_tbl_filt, otu_prevalence = otu_prevalence) %>%
  filter(otu_prevalence >= 10) %>%
  dplyr:: select(-con_mean, -case_mean, -otu_prevalence) %>%
  filter(genus_id != "NA")

rm(genus_tbl_filt)

genusnames = genus_tbl$genus_id
```

```r
genus_tbl_t <- genus_tbl %>%
  pivot_longer(cols= -1) %>%
  pivot_wider(names_from = "genus_id",values_from = "value") %>%
```

```
  rename(ID = name) %>%
  left_join(pheno_tbl, by = "ID")  # %>%
  #dplyr::select(-`NA`)
```

Filter Data at Species level

```
species_counts = aggTax(msd16s, lvl = "species", out = "matrix", norm = F)
species_id <- rownames(species_counts)
species_counts_tbl <- bind_cols(species_id = species_id, species_counts %>% as_tibble
species_counts_tbl
```

```
# A tibble: 754 × 993
   species_id     `100259` `100262` `100267` `100274` `100275` `100277` `100291`
   <chr>             <dbl>    <dbl>    <dbl>    <dbl>    <dbl>    <dbl>    <dbl>
 1 [Leptotrichia…        0        0        0        0        0        0        0
 2 Abiotrophia d…        0        0        0        0        0        0        1
 3 Acetobacter i…        0        0        0        0        0        0        0
 4 Acetobacter s…        0        0        0        0        0        0        0
 5 Acetonema lon…        0        0        0        0        0        0        0
 6 Acidaminococc…        6        0        0       26        0       31        0
 7 Acidaminococc…        0        0        0        0        0        0        0
 8 Acidithiobaci…        0        0       79        0        0        0        0
 9 Acidithiobaci…        0        0        0        0        0        0        0
10 Acidovorax ca…        0        0        0        0        0        0        0
# ℹ 744 more rows
# ℹ 985 more variables: `100292` <dbl>, `100293` <dbl>, `100294` <dbl>,
#   `100298` <dbl>, `100303` <dbl>, `100317` <dbl>, `100320` <dbl>,
#   `100322` <dbl>, `100341` <dbl>, `100353` <dbl>, `100356` <dbl>,
#   `100361` <dbl>, `100365` <dbl>, `100395` <dbl>, `100401` <dbl>,
#   `100403` <dbl>, `100437` <dbl>, `100457` <dbl>, `100462` <dbl>,
#   `100470` <dbl>, `100476` <dbl>, `100489` <dbl>, `100491` <dbl>, …
```

```
con_mean <- species_counts_tbl %>%
  dplyr:: select(all_of(control_to_select)) %>%
  rowMeans()

case_mean <- species_counts_tbl %>%
  dplyr:: select(all_of(case_to_select)) %>%
  rowMeans()

species_tbl_filt <- bind_cols(species_counts_tbl,
                              con_mean = con_mean,
                              case_mean = case_mean) %>%
  filter(con_mean >= 12 | case_mean >= 12)
```

```
otu_prevalence <- species_tbl_filt %>%
  dplyr:: select(-species_id) %>%
  mutate(across(everything(), ~ as.integer(. > 0))) %>%
  rowSums()
```

```
species_tbl <- bind_cols(species_tbl_filt, otu_prevalence = otu_prevalence) %>%
  filter(otu_prevalence >= 10) %>%
  dplyr:: select(-con_mean, -case_mean, -otu_prevalence)

rm(species_tbl_filt)
speciesnames = species_tbl$species_id[-1*which(species_tbl$species_id == "NA")]
```

```
species_tbl_t <- species_tbl %>%
  pivot_longer(cols= -1) %>%
  pivot_wider(names_from = "species_id",values_from = "value") %>%
  rename(ID = name) %>%
  left_join(pheno_tbl, by = "ID") %>%
  dplyr:: select(-`NA`)
```

Q1.1

```
compute_qaic <- function(model) {
  theta <- sum(residuals(model, type = "pearson")^2) / model$df.residual
  loglik <- logLik(update(model, family = poisson))
  qaic <- -2 * as.numeric(loglik) + 2 * model$rank * theta
  return(qaic)
}
```

```
results_genus_qpoisson <-
  map_df(genusnames, function(genusnames) {
    model <- glm(as.formula(paste(genusnames, "~ Type")),
                 data = genus_tbl_t,
                 family = quasipoisson)
    qaic <- compute_qaic(model)
    tidy(model, conf.int = TRUE) %>%
      # Add a column for the response variable
      mutate(response_variable = genusnames, qaic = qaic)
    }) %>%
  filter(term == "TypeControl") %>%
  arrange(p.value)
```

```
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
```

```
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
```

```
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
```

```r
result_qppoisson <- results_genus_qpoisson %>%
  filter(p.value < 0.05/41) %>%
  dplyr::select(response_variable, p.value, qaic) %>%
  print(n = Inf)
```

```
# A tibble: 9 × 3
  response_variable  p.value    qaic
  <chr>                <dbl>   <dbl>
1 Escherichia       1.20e-21 577156.
2 Prevotella        3.48e-18 577033.
```

```
3 Streptococcus     2.59e-10 576417.
4 Haemophilus       3.64e- 8 572044.
5 Bacteroides       1.25e- 6 576385.
6 Eubacterium       2.97e- 6 571740.
7 Lactobacillus     1.09e- 5 577251.
8 Campylobacter     1.52e- 4 573464.
9 Collinsella       1.13e- 3 572478.
```

```
count_qpposson <- result_qppoisson %>%
  summarise(count = n())

count_qpposson <- count_qpposson$count
```

```
library(MASS)
```

```
Attaching package: 'MASS'

The following object is masked from 'package:dplyr':

    select
```

```
results_genus_nb2 <-
  map_df(genusnames, function(genusnames) {
  #  print(response)
    model <- glm(as.formula(paste(genusnames, "~ Type")),
                 family = negative.binomial(20),
                 data = genus_tbl_t)
    qaic <- compute_qaic(model)
    tidy(model, conf.int = TRUE) %>%
      # Add a column for the response variable
      mutate(response_variable = genusnames, qaic = qaic)
    }) %>%
  filter(term == "TypeControl") %>%
  arrange(p.value)
```

```
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
```

```
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
```

```
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
```

```r
result_nb2 <- results_genus_nb2 %>%
  filter(p.value < 0.05/41) %>%
  dplyr::select(response_variable, p.value, qaic) %>%
  print(n = Inf)
```

```
# A tibble: 11 × 3
   response_variable  p.value    qaic
   <chr>                <dbl>   <dbl>
 1 Escherichia       1.95e-26 571190.
 2 Prevotella        1.88e-18 571098.
 3 Streptococcus     1.01e-11 571273.
```

```
 4 Haemophilus       9.70e-11 571402.
 5 Bacteroides       6.85e- 7 571557.
 6 Eubacterium       7.36e- 7 571272.
 7 Lactobacillus     2.71e- 6 571416.
 8 Campylobacter     3.82e- 5 572017.
 9 Collinsella       8.86e- 4 571550.
10 Weissella         1.16e- 3 572152.
11 Blautia           1.20e- 3 571600.
```

```r
count_nb2 <- result_nb2 %>%
  summarise(count = n())

count_nb2 <- count_nb2$count
```

```r
tidy_zeroinfl <- function(model) {
  coefs <- summary(model)$coefficients
  count_coefs <- as.data.frame(coefs$count)
  zero_coefs <- as.data.frame(coefs$zero)

  tidy_count <- count_coefs %>%
    rownames_to_column(var = "term") %>%
    mutate(component = "count")

  tidy_zero <- zero_coefs %>%
    rownames_to_column(var = "term") %>%
    mutate(component = "zero")

  tidy_model <- bind_rows(tidy_count, tidy_zero)
  names(tidy_model) <- c("term", "estimate", "std.error", "statistic", "p.value", "co

  return(tidy_model)
}
```

```r
compute_qaic_zeroinfl <- function(model) {
  # Calculate the theta (dispersion parameter)
  theta <- sum(residuals(model, type = "pearson")^2) / model$df.residual
  loglik <- logLik(model)
  k <- length(coef(model)) + length(model$coefficients$zero)  # Number of parameters
  qaic <- -2 * as.numeric(loglik) + 2 * k * theta
  return(qaic)
}
```

```r
results_genus_zfpoisson <-
  map_df(genusnames, function(genusnames) {
  #  print(response)
    model <- zeroinfl(as.formula(paste(genusnames, "~ Type")),
                  data = genus_tbl_t,
                  dist = "poisson")
    qaic <- compute_qaic_zeroinfl(model)
    tidy_zeroinfl(model) %>%
      # Add a column for the response variable
```

```
      mutate(response_variable = genusnames, qaic = qaic)
  }) %>%
  filter(term == "TypeControl") %>%
  arrange(p.value)

result_zfpoisson <- results_genus_zfpoisson %>%
  filter(p.value < 0.05/41) %>%
  dplyr::select(response_variable, p.value, qaic)

count_zfpoisson <- result_zfpoisson %>%
  summarise(count = n())

count_zfpoisson <- count_zfpoisson$count
```

```
results_genus_zfnb <-
  map_df(genusnames, function(genusnames) {
  #  print(response)
    model <- zeroinfl(as.formula(paste(genusnames, "~ Type")),
                  data = genus_tbl_t,
                  dist = "negbin")
    qaic <- compute_qaic_zeroinfl(model)
    tidy_zeroinfl(model) %>%
      # Add a column for the response variable
      mutate(response_variable = genusnames, qaic = qaic)
    }) %>%
  filter(term == "TypeControl") %>%
  arrange(p.value)
```

```
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
```

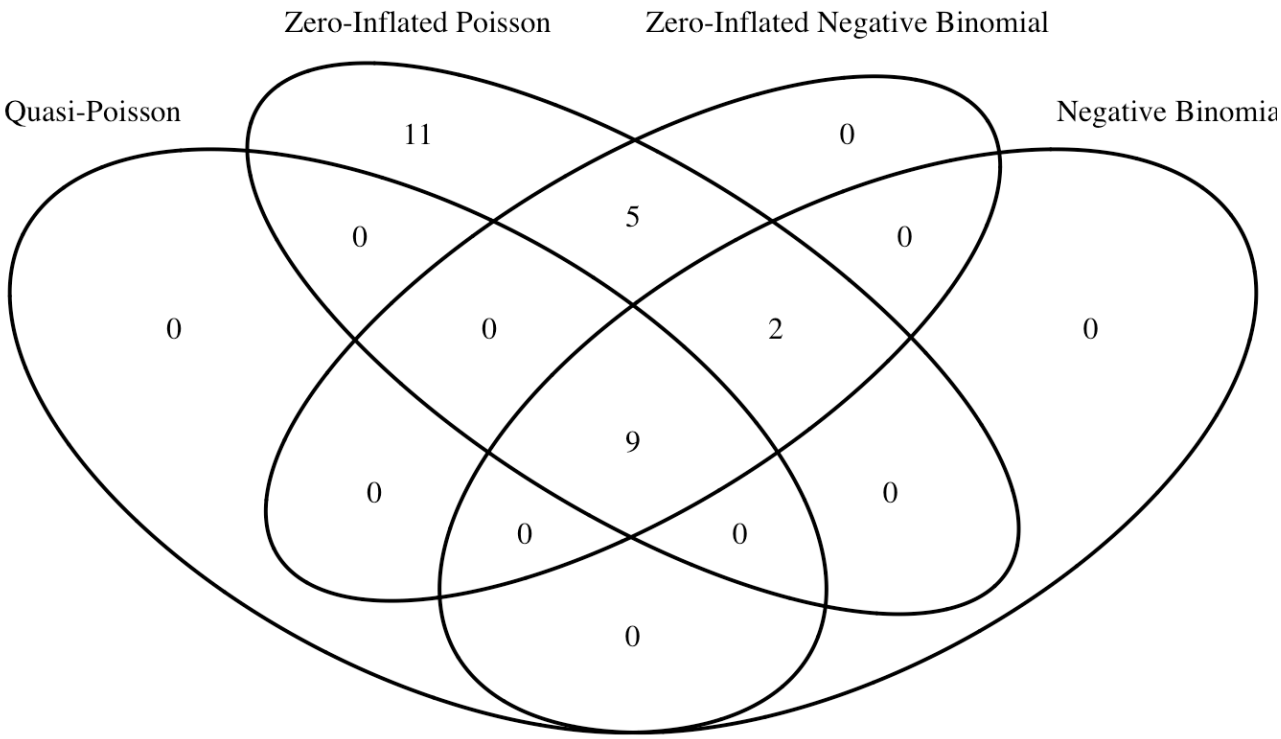
```
result_zfnb <- results_genus_zfnb %>%
  filter(p.value < 0.05/41) %>%
  dplyr::select(response_variable, p.value, qaic)

count_zfnb <- result_zfnb %>%
  summarise(count = n())

count_zfnb <- count_zfnb$count
```

```
tibble::tibble(
  Model = c("Quasi-Poisson", "Negative Binomial", "Zero Inflated Poisson", "Zero Infl
  Significant = c(count_qpposson, count_nb2, count_zfpoisson, count_zfnb)
)
```

```
# A tibble: 4 × 2
  Model                          Significant
  <chr>                                <int>
1 Quasi-Poisson                            9
2 Negative Binomial                       11
```

```
3 Zero Inflated Poisson                 39
4 Zero Inflated Negative Binomial       16
```

```r
list_response <- list(
  "Quasi-Poisson" = result_qppoisson$response_variable,
  "Negative Binomial" = result_nb2$response_variable,
  "Zero-Inflated Poisson" = result_zfpoisson$response_variable,
  "Zero-Inflated Negative Binomial" = result_zfnb$response_variable
)
grid.newpage()
# Generate the Venn diagram
venn.plot <- venn.diagram(
  x = list_response,
  category.names = c("Quasi-Poisson", "Negative Binomial", "Zero-Inflated Poisson", "
  filename = NULL,
  output = TRUE
)

# Display the Venn diagram
grid.draw(venn.plot)
```



Species Level

```
species_names <- names(species_tbl_t)[2:49]
```

```
results_species_qpoisson <-
  map_df(species_names, function(species_names) {
    model <- glm(as.formula(paste("`", species_names, "`", "~ Type", sep = "")),
                 data = species_tbl_t,
                 family = quasipoisson)
    qaic <- compute_qaic(model)
    tidy(model, conf.int = TRUE) %>%
      # Add a column for the response variable
      mutate(response_variable = species_names, qaic)
    }) %>%
  filter(term == "TypeControl") %>%
  arrange(p.value)
```

```
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
```

```
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
```

```
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
```

```
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
```

```
result_species_qpoisson <- results_species_qpoisson %>%
  filter(p.value < 0.05/49) %>%
  dplyr::select(response_variable, p.value, qaic)
```

```
count_species_qpoisson <- result_species_qpoisson %>%
  summarise(count = n())

count_species_qpoisson <- count_species_qpoisson$count
```

```
results_species_nb2 <-
  map_df(species_names, function(species_names) {
    model <- glm(as.formula(paste("`", species_names, "`", "~ Type", sep = "")),
                 data = species_tbl_t,
                 family = negative.binomial(20))
    qaic <- compute_qaic(model)
    tidy(model, conf.int = TRUE) %>%
      # Add a column for the response variable
      mutate(response_variable = species_names, qaic = qaic)
    }) %>%
  filter(term == "TypeControl") %>%
  arrange(p.value)
```

```
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
```

```
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
```

```
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
```

```
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
```

```
Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
  Consider formula(paste(x, collapse = " ")) instead.
```

```
result_species_nb2 <- results_species_nb2 %>%
  filter(p.value < 0.05/49) %>%
  dplyr::select(response_variable, p.value, qaic)



count_species_nb2 <- result_species_nb2 %>%
  summarise(count = n())

count_species_nb2 <- count_species_nb2$count
```

```
results_species_zfpoisson <-
  map_df(species_names, function(species_names) {
    model <- zeroinfl(as.formula(paste("`", species_names, "`", "~ Type", sep = "")),
                      data = species_tbl_t)
    qaic <- compute_qaic_zeroinfl(model)
    tidy_zeroinfl(model) %>%
      # Add a column for the response variable
      mutate(response_variable = species_names, qaic = qaic)
    }) %>%
  filter(term == "TypeControl") %>%
  arrange(p.value)

result_species_zfpoisson <- results_species_zfpoisson %>%
  filter(p.value < 0.05/49) %>%
  dplyr::select(response_variable, p.value, qaic)



count_species_zfpoisson <- result_species_zfpoisson %>%
  summarise(count = n())

count_species_zfpoisson <- count_species_zfpoisson$count
```

```
results_species_zfnb <-
  map_df(species_names, function(species_names) {
    model <- zeroinfl(as.formula(paste("`", species_names, "`", "~ Type", sep = "")),
                      dist = "negbin",
                      data = species_tbl_t)
    qaic <- compute_qaic_zeroinfl(model)
    tidy_zeroinfl(model) %>%
      # Add a column for the response variable
      mutate(response_variable = species_names, qaic = qaic)
    }) %>%
  filter(term == "TypeControl") %>%
  arrange(p.value)
```

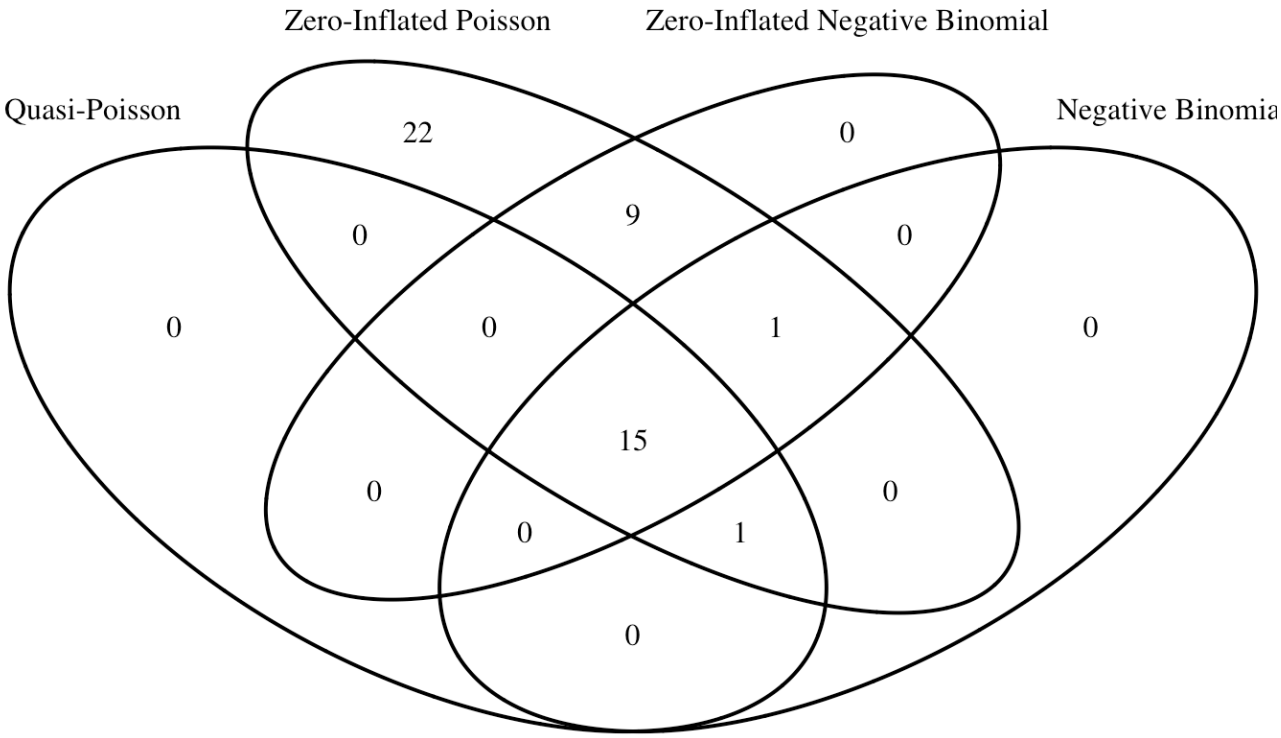
```
Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
```

```
result_species_zfnb <- results_species_zfnb %>%
  filter(p.value < 0.05/49) %>%
  dplyr::select(response_variable, p.value, qaic)



count_species_zfnb <- result_species_zfnb %>%
  summarise(count = n())

count_species_zfnb <- count_species_zfnb$count
```

```
tibble::tibble(
  Model = c("Quasi-Poisson", "Negative Binomial", "Zero Inflated Poisson", "Zero Infl
  Significant = c(count_species_qpoisson, count_species_nb2, count_species_zfpoisson,
)
```

```
# A tibble: 4 × 2
  Model                          Significant
  <chr>                                <int>
1 Quasi-Poisson                           16
2 Negative Binomial                       17
3 Zero Inflated Poisson                   67
4 Zero Inflated Negative Binomial         25
```

```
list_response <- list(
  "Quasi-Poisson" = result_species_qpoisson$response_variable,
  "Negative Binomial" = result_species_nb2$response_variable,
  "Zero-Inflated Poisson" = result_species_zfpoisson$response_variable,
  "Zero-Inflated Negative Binomial" = result_species_zfnb$response_variable
)
grid.newpage()
# Generate the Venn diagram
venn.plot <- venn.diagram(
  x = list_response,
  category.names = c("Quasi-Poisson", "Negative Binomial", "Zero-Inflated Poisson", "
  filename = NULL,
  output = TRUE
)

# Display the Venn diagram
grid.draw(venn.plot)
```



Q2

Genus Level

```
# Set the number of columns
num_columns <- 100
num_rows <- dim(genus_tbl_t)[1]
set.seed(10)

# Generate the tibble with 100 columns of random normal values
type1e <- function(genusname) {
  sim_tbl <- tibble(
    y = genus_tbl_t %>% dplyr::select(all_of(genusname)) %>% pull,
    as_tibble(
    matrix(sample(0:1, num_columns * num_rows, replace=T, prob=c(p_control, p_case)),
           nrow = num_rows, ncol = num_columns,
         dimnames = list(NULL, paste0("x", 1:num_columns)))))

  results <- map_df(names(sim_tbl)[-1],
                    ~ tidy(glm(reformulate(.x, response = "y"),
                               data = sim_tbl,
                               family = quasipoisson)),
                    .id = "variable") %>%
```

```
                filter(term != "(Intercept)")
  return(mean(results$p.value < 0.05))
}

type1e_tbl = tibble(genus_name = genusnames,
       etype_1_error = map_dbl(genusnames, type1e))
median_qpoisson_genus <- median(type1e_tbl$etype_1_error)
iqr_qpoisson_genus <- IQR(type1e_tbl$etype_1_error, na.rm = TRUE)
```

```
# Set the number of columns
num_columns <- 100
num_rows <- dim(genus_tbl_t)[1]
set.seed(10)

# Generate the tibble with 100 columns of random normal values
type1e <- function(genusname) {
  sim_tbl <- tibble(
    y = genus_tbl_t %>% dplyr::select(all_of(genusname)) %>% pull,
    as_tibble(
    matrix(sample(0:1, num_columns * num_rows, replace=T, prob=c(p_control, p_case)),
           nrow = num_rows, ncol = num_columns,
         dimnames = list(NULL, paste0("x", 1:num_columns)))))

  results <- map_df(names(sim_tbl)[-1],
                ~ tidy(glm(reformulate(.x, response = "y"),
                           data = sim_tbl,
                           family = negative.binomial(20),
                           control = glm.control(maxit = 100))),
                .id = "variable") %>%
              filter(term != "(Intercept)")
  return(mean(results$p.value < 0.05))
}

type1e_tbl = tibble(genus_name = genusnames,
       etype_1_error = map_dbl(genusnames, type1e))

median_nb_genus <- median(type1e_tbl$etype_1_error)
iqr_nb_genus <- IQR(type1e_tbl$etype_1_error, na.rm = TRUE)
```

```
# Set the number of columns
num_columns <- 100
num_rows <- dim(genus_tbl_t)[1]
set.seed(10)

# Generate the tibble with 100 columns of random normal values
type1e <- function(genusname) {
  sim_tbl <- tibble(
    y = genus_tbl_t %>% dplyr::select(all_of(genusname)) %>% pull,
    as_tibble(
    matrix(sample(0:1, num_columns * num_rows, replace=T, prob=c(p_control, p_case)),
           nrow = num_rows, ncol = num_columns,
         dimnames = list(NULL, paste0("x", 1:num_columns)))))
```

```
  results <- map_df(names(sim_tbl)[-1],
                    ~ tidy_zeroinfl(zeroinfl(reformulate(.x, response = "y"),
                                             data = sim_tbl,
                                             control = zeroinfl.control(maxit = 100))),
                    .id = "variable") %>%
                  filter(term != "(Intercept)")
  return(mean(results$p.value < 0.05))
}

type1e_tbl = tibble(genus_name = genusnames,
       etype_1_error = map_dbl(genusnames, type1e))

median_zfpoisson_genus <- median(type1e_tbl$etype_1_error)
iqr_zfpoisson_genus <- IQR(type1e_tbl$etype_1_error, na.rm = TRUE)
```

```
# Set the number of columns
num_columns <- 100
num_rows <- dim(genus_tbl_t)[1]
set.seed(10)

# Generate the tibble with 100 columns of random normal values
type1e <- function(genusname) {
  sim_tbl <- tibble(
    y = genus_tbl_t %>% dplyr::select(all_of(genusname)) %>% pull,
    as_tibble(
    matrix(sample(0:1, num_columns * num_rows, replace=T, prob=c(p_control, p_case)),
           nrow = num_rows, ncol = num_columns,
         dimnames = list(NULL, paste0("x", 1:num_columns)))))

  results <- map_df(names(sim_tbl)[-1],
                    ~ tidy_zeroinfl(zeroinfl(reformulate(.x, response = "y"),
                                             data = sim_tbl,
                                             dist = "negbin",
                                             control = zeroinfl.control(maxit = 100))),
                    .id = "variable") %>%
                  filter(term != "(Intercept)")
  return(mean(results$p.value <0.05, na.rm = TRUE))
}

type1e_tbl = tibble(genus_name = genusnames,
       etype_1_error = map_dbl(genusnames, type1e))
```

```
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
```

```
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
```

```
Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 8.59577e-33FALSE

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
```

```
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
```

```
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
```

```
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
```

```
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
```

```r
median_zfnb_genus <- median(type1e_tbl$etype_1_error)
iqr_zfnb_genus <- IQR(type1e_tbl$etype_1_error, na.rm = TRUE)
```

Species Level

```r
# Set the number of columns
num_columns <- 100
num_rows <- 992
set.seed(10)

# Generate the tibble with 100 columns of random normal values
type1e <- function(speciesnames) {
  sim_tbl <- tibble(
    y = species_tbl_t %>% dplyr::select(all_of(speciesnames)) %>% pull,
```

```
    as_tibble(
    matrix(sample(0:1, num_columns * num_rows, replace=T, prob=c(p_control, p_case)),
          nrow = num_rows, ncol = num_columns,
         dimnames = list(NULL, paste0("x", 1:num_columns)))))

  results <- map_df(names(sim_tbl)[-1],
                    ~ tidy(glm(reformulate(.x, response = "y"),
                               data = sim_tbl), family = quasipoisson),
                    .id = "variable") %>%
                filter(term != "(Intercept)")
  return(mean(results$p.value < 0.05))
}

type1e_tbl = tibble(
       species_name = speciesnames,
       etype_1_error = map_dbl(speciesnames, type1e))
median_gpoisson_species <- median(type1e_tbl$etype_1_error)
iqr_qpoisson_species <- IQR(type1e_tbl$etype_1_error, na.rm = TRUE)
```

```
# Set the number of columns
num_columns <- 100
num_rows <- 992


# Generate the tibble with 100 columns of random normal values
type1e <- function(speciesname) {
  sim_tbl <- tibble(
    y = species_tbl_t %>%
      dplyr::select(all_of(speciesname)) %>% pull,
    as_tibble(
    matrix(sample(0:1, num_columns * num_rows, replace=T, prob=c(p_control, p_case)),
           nrow = num_rows, ncol = num_columns,
           dimnames = list(NULL, paste0("x", 1:num_columns)))))

    results <- map_df(names(sim_tbl)[-1],
                ~ tidy(glm(reformulate(.x, response = "y"),
                           data = sim_tbl,
                           family = negative.binomial(30),
                           control = glm.control(maxit = 100))),
                .id = "variable") %>%
              filter(term != "(Intercept)")
  return(mean(results$p.value < 0.05))
}

type1e_tbl = tibble(
       species_name = speciesnames,
       etype_1_error = map_dbl(speciesnames, type1e))
median_nb_species <- median(type1e_tbl$etype_1_error)
iqr_nb_species <- IQR(type1e_tbl$etype_1_error, na.rm = TRUE)
```

```
# Set the number of columns
num_columns <- 100
```

```
num_rows <- 992

# Generate the tibble with 100 columns of random normal values
type1e <- function(speciesnames) {
  sim_tbl <- tibble(
    y = species_tbl_t %>% dplyr::select(all_of(speciesnames)) %>% pull,
    as_tibble(
    matrix(sample(0:1, num_columns * num_rows, replace=T, prob=c(p_control, p_case)),
          nrow = num_rows, ncol = num_columns,
         dimnames = list(NULL, paste0("x", 1:num_columns)))))

  results <- map_df(names(sim_tbl)[-1],
                ~ tidy_zeroinfl(zeroinfl(reformulate(.x, response = "y"),
                                         data = sim_tbl)),
                .id = "variable") %>%
             filter(term != "(Intercept)")
  return(mean(results$p.value < 0.05))
}

type1e_tbl = tibble(
       species_name = speciesnames,
       etype_1_error = map_dbl(speciesnames, type1e))
median_zfpoisson_species <- median(type1e_tbl$etype_1_error)
iqr_zfpoisson_species <- IQR(type1e_tbl$etype_1_error, na.rm = TRUE)
```

```
# Set the number of columns
num_columns <- 100
num_rows <- 992

# Generate the tibble with 100 columns of random normal values
type1e <- function(speciesnames) {
  sim_tbl <- tibble(
    y = species_tbl_t %>% dplyr::select(all_of(speciesnames)) %>% pull,
    as_tibble(
    matrix(sample(0:1, num_columns * num_rows, replace=T, prob=c(p_control, p_case)),
          nrow = num_rows, ncol = num_columns,
         dimnames = list(NULL, paste0("x", 1:num_columns)))))
  results <- map_df(names(sim_tbl)[-1],
                ~ tidy_zeroinfl(zeroinfl(reformulate(.x, response = "y"),
                                         data = sim_tbl,
                                         dist = "negbin",
                                         control = zeroinfl.control(maxit = 100))),
                .id = "variable") %>%
             filter(term != "(Intercept)")

  return(mean(results$p.value <0.05, na.rm = TRUE))
}

type1e_tbl = tibble(
       species_name = speciesnames,
       etype_1_error = map_dbl(speciesnames, type1e))
```

```
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 6.72602e-33FALSE

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 4.334e-33FALSE

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 5.47762e-33FALSE

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 6.85781e-33FALSE

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 4.68351e-33FALSE

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 6.67376e-33FALSE

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 6.46229e-33FALSE

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 6.20379e-33FALSE

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 6.88214e-33FALSE

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 4.68402e-33FALSE
```

```
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 4.62637e-33FALSE

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 4.5993e-33FALSE

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 4.15884e-33FALSE

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 6.26036e-33FALSE

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 4.36884e-33FALSE

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 4.40746e-33FALSE

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 4.20719e-33FALSE

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 3.93694e-33FALSE

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
```

```
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
```

```
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
```

```
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 3.48154e-33FALSE
```

```
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 6.61912e-33FALSE

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 4.185e-33FALSE

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 6.96499e-33FALSE

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 3.54836e-33FALSE

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 7.24416e-33FALSE

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 3.81303e-33FALSE

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
```

```
"negbin", : optimization failed to converge
Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 3.86175e-33FALSE

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 7.34914e-33FALSE

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced
```

```
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
```

```
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
```

```
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
```

```
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
```

```
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
```

```
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced
```

```
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
```

```
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
```

```
Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in value[[3L]](cond): system is computationally singular: reciprocal
condition number = 6.58627e-33FALSE

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
```

```
Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
```

```
Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in sqrt(diag(vc)[np]): NaNs produced

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced

Warning in zeroinfl(reformulate(.x, response = "y"), data = sim_tbl, dist =
"negbin", : optimization failed to converge

Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
Warning in sqrt(diag(object$vcov)): NaNs produced
```

```r
median_zfnb_species <- median(type1e_tbl$etype_1_error)
iqr_zfnb_species <- IQR(type1e_tbl$etype_1_error, na.rm = TRUE)
```

```r
result_tbl <- tibble(
  Method = c("Quasi-Poisson", "Negative Binomial", "Zero Inflated Poisson", "Zero Inf
  Genus = c(
    paste0(median_qpoisson_genus, " (IQR: ", iqr_qpoisson_genus, ")"),
    paste0(median_nb_genus, " (IQR: ", iqr_nb_genus, ")"),
    paste0(median_zfpoisson_genus, " (IQR: ", iqr_zfpoisson_genus, ")"),
    paste0(median_zfnb_genus, " (IQR: ", iqr_zfnb_genus, ")")
  ),
  Species = c(
    paste0(median_gpoisson_species, " (IQR: ", iqr_qpoisson_species, ")"),
    paste0(median_nb_species, " (IQR: ", iqr_nb_species, ")"),
    paste0(median_zfpoisson_species, " (IQR: ", iqr_zfpoisson_species, ")"),
    paste0(median_zfnb_species, " (IQR: ", iqr_zfnb_species, ")")
  )
)
```

```
        # Print the result tibble
        print(result_tbl)
```

```
# A tibble: 4 × 3
  Method                          Genus                                  Species
  <chr>                           <chr>                                  <chr>
1 Quasi-Poisson                   0.05 (IQR: 0.04)                       0.04 (…
2 Negative Binomial               0.07 (IQR: 0.05)                       0.06 (…
3 Zero Inflated Poisson           0.495 (IQR: 0.03)                      0.4925…
4 Zero Inflated Negative Binomial 0.414715719063545 (IQR: 0.08971075998… 0.3959…
```

Q3

```
        mean(result_species_qpoisson$qaic)
```

```
[1] 392699.5
```

```
        mean(result_species_nb2$qaic)
```

```
[1] 390427
```

```
        mean(result_species_zfpoisson$qaic)
```

```
[1] 136633.8
```

```
        mean(result_species_zfnb$qaic)
```

```
[1] 6036.67
```

```
        mean(result_qppoisson$qaic)
```

```
[1] 574885.3
```

```
        mean(result_nb2$qaic)
```

```
[1] 571502.4
```

```
        mean(result_zfpoisson$qaic)
```

```
[1] 225762.9
```

```
        mean(result_zfnb$qaic)
```

```
[1] 7109.753
```

By observing the QAIC(A similar measure of goodness of fit as AIC), the zero inflated negative bionomial model seems to be the best fit to model the data.

By observing the type I error, we can see that the quasi poisson model has the smallest type I error. The zero inflated poisson model has the highest type I error. The zero inflated negative binomial model also has very high type I error. The negative binomial model has small type I error closed to quasi binomial model. We can conclude that zero-inflated model inflates the type I error.

Based on these information, I would conclude that using quasi-binomial model is the most ideal case among the four models. Although it has the highest QAIC, it has the smallest type I error.

To summarize the result of quasi-binomial model, we have QAIC = 392699.5 for Genus and QAIC = 574885.3 for Species. The type I error is 0.05 for Genus and 0.04 for Species.

Redo Midterm question

Question 2

1

(1) Answer:

The deviance of fitted model = 0.95443=residual deviance

It is calculated by $2\frac{\log L_\Omega}{\log L_\omega}$

(2) Answer:

H_0: 2 models are not distinguishable $(\beta_1 = \beta_2 = \beta_3 = 0)$

H_A: 2 models are distinguishable (at least one $\beta_i \neq 0$)

When df = 4, the p value for critical value 0.95443 is between 0.1 and 0.9 which is greater than 0.05. Hence, we do not have enough evidence to conclude fitted model and saturated model are significantly different. Hence, the fitted model is a good fit to the data.

(3) Answer:

$D_{null} - D_w = 21.33202 - 0.95443 = 20.37759$

$Df_{null} - Df_w = 7 - 4 = 3$

According to the table of Chi-square distribution, the p-value is less than 0.001.

Hence, we conclude 3 predictors provide a significant better fit than the intercept only model.

(4) Answer:

$\log(\frac{p_1}{1-p_i}) = -2.6163 + 0.6223sex_i + 0.6157age_i + 0.362ecg_i$

$log(odd) = -2.6163 + 0.6223 + 0.6157 + 0.362 = -1.0163$

$odd = e^{-1.0163} \approx 0.3619$

(5) Answer:

8 parameters in model. The deviance will be 0 since when number of parameter is equivalent to number of observation, we have a saturated model which has deviance 0.

(6) Answer:

Hessian is $\sum_i [-(x_i^T)^T(\frac{1}{(1+e^{x_i^T\beta})^2})(e^{x_i^T\beta})(x_i^T)]$

For all vectors a,

$$\sum_i [a^T(x_i^T)^T(\frac{1}{(1+e^{x_i^T\beta})^2})(e^{x_i^T\beta})(x_i^T)a] = \sum_i [(x_i^Ta)^T(\frac{e^{x_i^T\beta}}{(1+e^{x_i^T\beta})^2})(x_i^Ta)] = \sum_i [(x_i^Ta)^2(\frac{e^{x_i^T\beta}}{(1+e^{x_i^T\beta})^2})$$

Since exponential function is always positive, the Hessian matrix is positive semidefinite by Energy-based definition. Therefore, the log-likelihood function is concave.

(7) Answer:

It will be the same as binomial model since the extended bernoulli model is equivalent to the binomial mode.

Question 8

(1) Answer:

3+4=7 parameters in the model

(2) Answer:

$\frac{1}{exp(\theta_1 - 40\beta_{age} - \beta_{gender} - \beta_{smoke})}$

(3) Answer:

$exp(-\beta_{smoke} - \beta_{npt})$

(4) Answer:

$exp(\beta_{smoke} + \beta_{npt})$

(5) Answer:

$exp(-\beta_{smoke} - \beta_{npt})$