Experiment 1–Flight Cylinder Data

Analysis of Flight Cylinder Data: Overview

Reads in data on how many out of 10 navel orangeworm adults left a flight cylinder within 24 hours, creates a summary, and performs a 2-way analysis of Deviance with moth performance (y/n, with y as number leaving and n a the number of moths in the flight cylinder).

Load data into the Global Environment (working memory)

```
library(readxl)       # Read Excel files
library(janitor)      # Clean variable names
df_fltcyl <- read_excel("Dataset.xlsx", sheet = "Flight_ability_assay")
df_fltcyl <- clean_names(df_fltcyl)
df_fltcyl
```

```
## # A tibble: 60 × 4
##      rep sex    dose twentyfour_out
##    <dbl> <chr> <dbl>          <dbl>
##  1     1 f         0              9
##  2     1 f       100              9
##  3     1 f       150              9
##  4     1 f       200              9
##  5     1 f       250              7
##  6     1 f       300             10
##  7     2 f         0              7
##  8     2 f       100              8
##  9     2 f       150              8
## 10     2 f       200              9
## # ℹ 50 more rows
```

Get and export treatment means and SE

Will use dplyr to get a table of sample size, mean, and standard error by sex and radiation dose.

```
library(dplyr)                  # Easy and efficient data wrangling and
summary
library(FSA)                    # Provides an se() function for standard error
tbl_escapes <- df_fltcyl %>%
  group_by(sex,dose) %>%
  summarise(nObs = n(),
            mn = mean(twentyfour_out),
            sem = se(twentyfour_out))
write.csv(tbl_escapes,"tbl01_flt_cyl_escapes.csv", row.names = F)
tbl_escapes
```

```
## # A tibble: 12 × 5
## # Groups:   sex [2]
##    sex    dose  nObs    mn   sem
##    <chr> <dbl> <int> <dbl> <dbl>
##  1 f         0     5   8.4 0.6
##  2 f       100     5   6.6 0.812
##  3 f       150     5   7.8 0.374
##  4 f       200     5   7.8 0.583
##  5 f       250     5   7.4 0.510
##  6 f       300     5   6.8 0.860
##  7 m         0     5   8.8 0.583
##  8 m       100     5   8   0.316
##  9 m       150     5   8.8 0.2
## 10 m       200     5   8.4 0.510
## 11 m       250     5   8.6 0.245
## 12 m       300     5   8.4 0.678
```

Statistical Analysis

Will perform the Analysis of Deviance described above using the R-native glm() function. The output provides general information on degrees of freedom and significance of the main factors and interaction.

```
# Identify Sex and Dose as Factors
df_fltcyl$sex <- as.factor(df_fltcyl$sex)
df_fltcyl$dose <- as.factor(df_fltcyl$dose)

# Contastant for number at risk
df_fltcyl$n <- 10
df_fltcyl <- df_fltcyl %>% rename(y = twentyfour_out)

# Create parameter for non-event
df_fltcyl <- df_fltcyl %>%
  mutate(not_y = n - y)

# Fit a logistic regression model using glm with binomial family
m1 <- glm(cbind(y,not_y) ~ sex * dose, data = df_fltcyl, family = binomial)

# Perform analysis of deviance using ANOVA for the GLM
anova_results <- anova(m1, test = "Chisq")

# Print results
print(anova_results)
```

```
## Analysis of Deviance Table
##
## Model: binomial, link: logit
##
## Response: cbind(y, not_y)
```

```
##
## Terms added sequentially (first to last)
##
##
##          Df Deviance Resid. Df Resid. Dev Pr(>Chi)
## NULL                        59     72.852
## sex       1  10.0391        58     62.813 0.001533 **
## dose      5   6.9998        53     55.814 0.220656
## sex:dose  5   0.9059        48     54.908 0.969792
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The 'car' package provides an alternative Anova() function that provides type II degrees of freedom. This is often seen as preferable.

```
library(car)
Anova(m1, type = "II")

## Analysis of Deviance Table (Type II tests)
##
## Response: cbind(y, not_y)
##          LR Chisq Df Pr(>Chisq)
## sex       10.1565  1   0.001438 **
## dose       6.9998  5   0.220656
## sex:dose   0.9059  5   0.969792
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Experiment 1–Longevity data

Analysis of longevity data

Done at the same time as the flight cylinder assay (previous script). Same-sex groups of 10 moths (all male or all female) in a cage were examined on a daily basis, and the number dead since the previous interval and the number of survivors were recorded. These data are summarized using Kaplan-Meier survivorship plots and Cox proportional hazard analysis. The data were recorded as groups, but the packages used here require one record per moth.

Since around a third of the moths examined were alive at the end of the observation period, a mean longevity has limited meaning while an estimate of the median longevity is arguably more informative. The package used to generate the survival curve is also used to get an estimate of median survival for each dose with a 95% confidence interval.

```
library(readxl)
library(dplyr)
df_lngvt <- read_excel("Dataset.xlsx", sheet = "nowcox")
df_lngvt <- df_lngvt %>%
  rename(day = dor,
         gy = dose)
df_lngvt
```

```
## # A tibble: 420 × 7
##      rep   day sex      gy  mort alive total
##    <dbl> <dbl> <chr> <dbl> <dbl> <dbl> <dbl>
##  1     1     1 f         0     0    10    10
##  2     1     2 f         0     1     9    10
##  3     1     3 f         0     1     8    10
##  4     1     4 f         0     0     8    10
##  5     1     5 f         0     2     6    10
##  6     1     6 f         0     3     3    10
##  7     1     7 f         0     3     0    10
##  8     2     1 f         0     0    10    10
##  9     2     2 f         0     0    10    10
## 10     2     3 f         0     2     8    10
## # ℹ 410 more rows
```

Recode from aggregated to case format

First using the dplyr and tidyr package to creat 1 row for each for death records

```
library(tidyr)

# For rows with deaths, create an observation per death using tidyr's uncount
()
events <- df_lngvt %>%
  filter(mort > 0) %>%                    # only rows where at least one dea
```

```
th occurred
  mutate(time = day, event = 1) %>%          # assign time and indicate event
= 1 (death)
  select(rep, day, sex, gy, time, mort, event) %>%
  uncount(weights = mort)                  # expand each row 'mort' times

# Preview event data:
events
```

```
## # A tibble: 429 × 6
##      rep   day sex      gy  time event
##    <dbl> <dbl> <chr> <dbl> <dbl> <dbl>
##  1     1     2 f         0     2     1
##  2     1     3 f         0     3     1
##  3     1     5 f         0     5     1
##  4     1     5 f         0     5     1
##  5     1     6 f         0     6     1
##  6     1     6 f         0     6     1
##  7     1     6 f         0     6     1
##  8     1     7 f         0     7     1
##  9     1     7 f         0     7     1
## 10     1     7 f         0     7     1
## # ℹ 419 more rows
```

All moths still alive on the last day of observation (day 7) are recorded as censored on that day.

```
censored <- df_lngvt %>%
  group_by(rep, sex, gy) %>%
  slice_tail(n = 1) %>%    # select the final monitoring interval for each gr
oup
  ungroup() %>%
  filter(alive > 0) %>%    # only if there are survivors
  mutate(time = day, event = 0) %>%  # assign time and indicate event = 0 (ce
nsored)
  select(rep, day, sex, gy, time, event, alive) %>%
  uncount(weights = alive) # expand each row 'alive' times

# Preview censored data:
censored
```

```
## # A tibble: 171 × 6
##      rep   day sex      gy  time event
##    <dbl> <dbl> <chr> <dbl> <dbl> <dbl>
##  1     1     7 f       100     7     0
##  2     1     7 f       100     7     0
##  3     1     7 f       150     7     0
##  4     1     7 f       250     7     0
##  5     1     7 f       250     7     0
##  6     1     7 f       250     7     0
```

```
##  7     1     7 f       250     7     0
##  8     1     7 m         0     7     0
##  9     1     7 m         0     7     0
## 10     1     7 m         0     7     0
## # ℹ 161 more rows
```

Now combine the events and censored observations

```
# Combine events and censored observations
indiv_data <- bind_rows(events, censored)

# (Optional) Order the data, if desired:
indiv_data <- indiv_data %>% arrange(rep, sex, gy, time)

# Make Sex and Dose (Gy) explicitly factors
indiv_data$sex <- as.factor(indiv_data$sex)
indiv_data$gy <- as.factor(indiv_data$gy)

# One last look
indiv_data
```

```
## # A tibble: 600 × 6
##       rep   day sex   gy     time event
##     <dbl> <dbl> <fct> <fct> <dbl> <dbl>
##  1     1     2 f     0         2     1
##  2     1     3 f     0         3     1
##  3     1     5 f     0         5     1
##  4     1     5 f     0         5     1
##  5     1     6 f     0         6     1
##  6     1     6 f     0         6     1
##  7     1     6 f     0         6     1
##  8     1     7 f     0         7     1
##  9     1     7 f     0         7     1
## 10     1     7 f     0         7     1
## # ℹ 590 more rows
```

Note here that event = 1 mean “event happened” (moth died), and event = 0 means moth was last seen alive

Use the Cox Proportional Hazards model to evaluate impact

The R package ‘Survival’ will be used to implement the Cox model and evaluate the impact of sex and irradiation

```
library(survival)

# Create the survival object
surv_obj <- with(indiv_data, Surv(time = time, event = event))

# Fit the Cox model (adjust predictors as appropriate)
```

```r
cox_fit <- coxph(surv_obj ~ sex + gy, data = indiv_data)

# View parameter estimates for the printed model
summary(cox_fit)
```

```
## Call:
## coxph(formula = surv_obj ~ sex + gy, data = indiv_data)
##
##   n= 600, number of events= 429
##
##            coef exp(coef)  se(coef)      z Pr(>|z|)
## sexm  -0.183926  0.831997  0.096956 -1.897   0.0578 .
## gy100  0.081795  1.085233  0.168802  0.485   0.6280
## gy150 -0.008396  0.991639  0.170342 -0.049   0.9607
## gy200  0.093471  1.097978  0.169682  0.551   0.5817
## gy250  0.145092  1.156146  0.169137  0.858   0.3910
## gy300  0.352740  1.422961  0.163446  2.158   0.0309 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##       exp(coef) exp(-coef) lower .95 upper .95
## sexm     0.8320     1.2019    0.6880     1.006
## gy100    1.0852     0.9215    0.7795     1.511
## gy150    0.9916     1.0084    0.7102     1.385
## gy200    1.0980     0.9108    0.7873     1.531
## gy250    1.1561     0.8649    0.8299     1.611
## gy300    1.4230     0.7028    1.0329     1.960
##
## Concordance= 0.555  (se = 0.016 )
## Likelihood ratio test= 10.3  on 6 df,   p=0.1
## Wald test            = 10.57  on 6 df,   p=0.1
## Score (logrank) test = 10.63  on 6 df,   p=0.1
```

The Likelihood ratio and the Wald test provide similar results, but the former would be considered more robust for small data sets while the latter would be considered more informative for large data sets. The P value of 0.1 and the concordance of 0.555 indicate that the model has not detected important effects. We can also examine the model using a type II ANOVA.

```r
library(car)
print(Anova(cox_fit))
```

```
## Analysis of Deviance Table (Type II tests)
##     LR Chisq Df Pr(>Chisq)
## sex   3.6044  1    0.05763 .
## gy    6.4243  5    0.26709
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The type II analysis provides a simple and robust parameter estimates indicating “not significant” (or an almost significant effect of sex, if one wishes to make that argument).

Another way to examine the current model is to generate estimates of median longevity. When the sexes are examined separately, the upper confidence cannot be estimated in some cases.

```
sfit_2way <- survfit(surv_obj ~ sex + gy, data = indiv_data)
print(sfit_2way)
```

```
## Call: survfit(formula = surv_obj ~ sex + gy, data = indiv_data)
##
##                 n events median 0.95LCL 0.95UCL
## sex=f, gy=0    50     43      6       6       7
## sex=f, gy=100 50     33      7       6      NA
## sex=f, gy=150 50     36      6       6       7
## sex=f, gy=200 50     38      6       5       7
## sex=f, gy=250 50     34      6       5       7
## sex=f, gy=300 50     41      6       5       7
## sex=m, gy=0    50     27      7       7      NA
## sex=m, gy=100 50     38      6       6       7
## sex=m, gy=150 50     32      7       6      NA
## sex=m, gy=200 50     31      7       6      NA
## sex=m, gy=250 50     36      6       6       7
## sex=m, gy=300 50     40      6       5       7
```

Examination of irradiation only (sexes pooled)

The primary purpose here is to get a more robust estimate of the 95% confidence interval for the median estimates.

```
cox_fit_dose_only <- coxph(surv_obj ~ gy, data = indiv_data)
Anova(cox_fit_dose_only)
```

```
## Analysis of Deviance Table
##  Cox model: response is surv_obj
## Terms added sequentially (first to last)
##
##       loglik  Chisq Df Pr(>|Chi|)
## NULL -2530.6
## gy   -2527.2 6.6951  5     0.2443
```

Still not significant, in case anyone wondered.

```
km_fit <- survfit(surv_obj ~ gy, data = indiv_data)
print(km_fit)
```

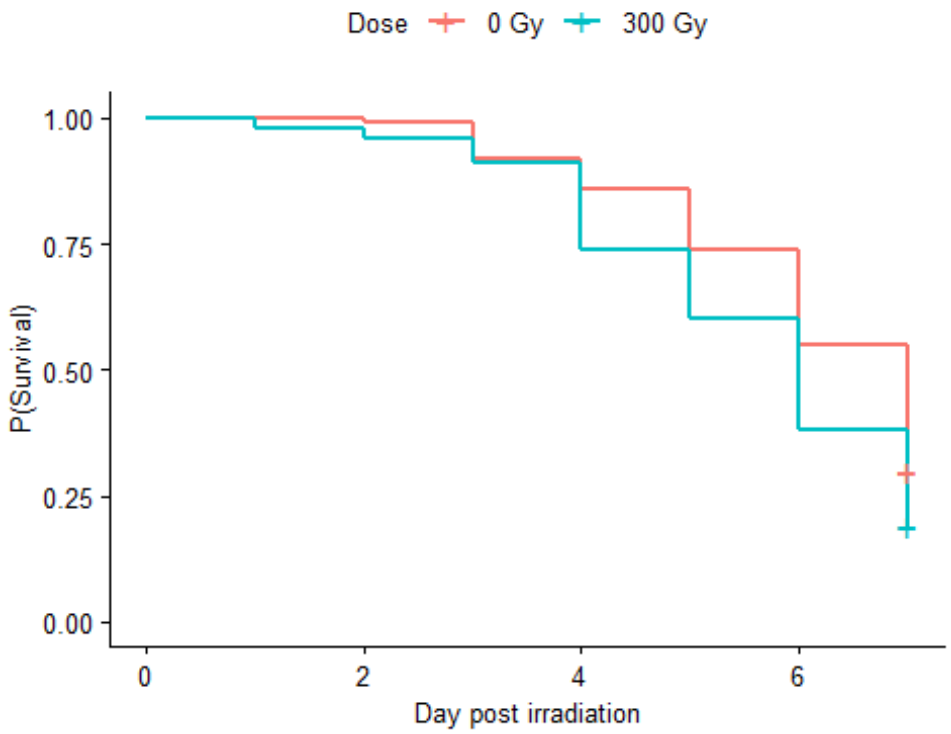
```
## Call: survfit(formula = surv_obj ~ gy, data = indiv_data)
##
##          n events median 0.95LCL 0.95UCL
## gy=0   100     70      7       6       7
```

```
## gy=100 100     71      6       6       7
## gy=150 100     68      7       6       7
## gy=200 100     69      6       6       7
## gy=250 100     70      6       6       7
## gy=300 100     81      6       6       6
```

Using the pooled sexes provides a larger sample and allows an upper 95% CL for all doses, so this is probably better for reporting.

For visualization, presenting survival curves for only the 0 and 300 Gy levels seems to offer a cleaner plot.

```
library(survminer)                    # Required for surv_fit() function
# Create set with only 0 and 300 Gy
indiv_data2 <- indiv_data %>%
  filter(gy == 0 | gy == 300)
# Drop now unused levels
indiv_data2$gy <- droplevels(indiv_data2$gy)

# Create another survival object
surv_obj2 <- with(indiv_data2, Surv(time = time, event = event))

# Fit Kaplan-Meier Curve
km_fit2 <- surv_fit(surv_obj2 ~ gy, data = indiv_data2)

# Plot survival curves
p1 <- ggsurvplot(km_fit2, data = indiv_data2, pval = FALSE,
                 legend.title = "Dose", legend.labs = c("0 Gy","300 Gy"))

# Format the image to be saved
p1$plot <- p1$plot +
  xlab("Day post irradiation") +
  ylab("P(Survival)") +
  theme(axis.text.x = element_text(color = "black", size = 10),
        axis.text.y = element_text(color = "black", size = 10),
        axis.title.x = element_text(color = "black", size = 10),
        axis.title.y = element_text(color = "black", size = 10),
        legend.title = element_text(color = "black", size = 10),
        legend.text = element_text(color = "black", size = 10))

# Save the image to local storage, then display the image
ggsave(filename = "survival_curve.jpg", plot = p1$plot,
       width = 2.83, height = 1.89, dpi = 300)

p1
```

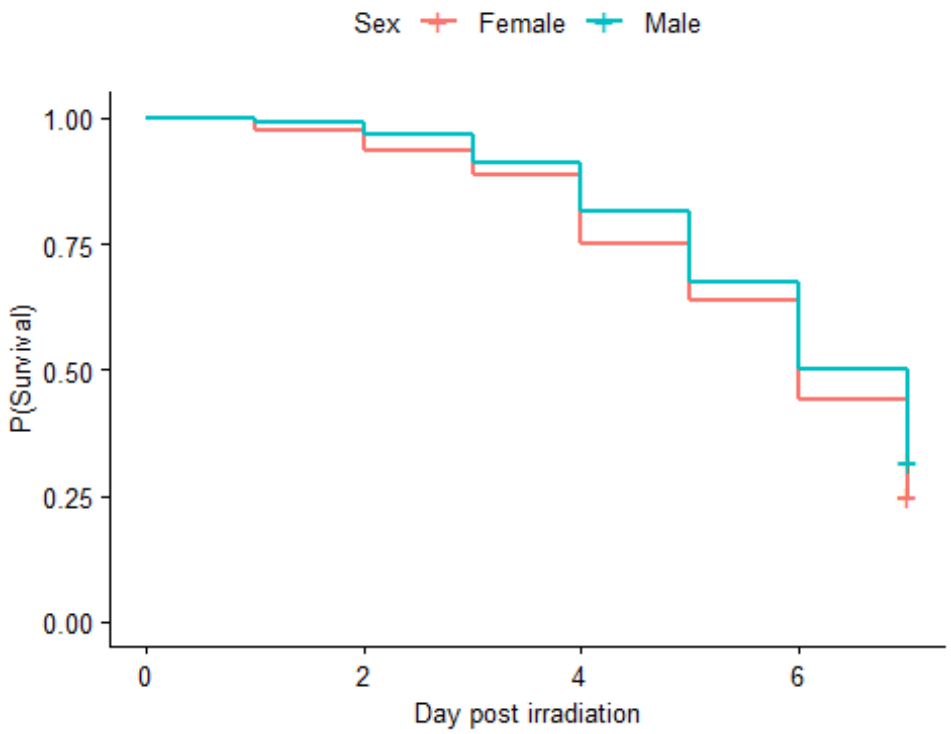


Examination of sex only (x-ray doses pooled)

In case we want to visualize the data in this manner, we can repeat the process in the previous section

```
# Since there are only two sexes, the earlier data set and surv_obj will work

# Fit Kaplan-Meier Curve
km_fit3 <- surv_fit(surv_obj ~ sex, data = indiv_data)

# Plot survival curves
p2 <- ggsurvplot(km_fit3, data = indiv_data2, pval = FALSE,
                 legend.title = "Sex", legend.labs = c("Female","Male"))

# Format the image to be saved
p2$plot <- p2$plot +
  xlab("Day post irradiation") +
  ylab("P(Survival)") +
  theme(axis.text.x = element_text(color = "black", size = 10),
        axis.text.y = element_text(color = "black", size = 10),
        axis.title.x = element_text(color = "black", size = 10),
        axis.title.y = element_text(color = "black", size = 10),
        legend.title = element_text(color = "black", size = 10),
        legend.text = element_text(color = "black", size = 10))

# Save the image to local storage, then display the image
```

```
ggsave(filename = "survival_curve_sex.jpg", plot = p2$plot,
       width = 2.83, height = 1.89, dpi = 300)

p2
```



Experiment 2 – Mating

Overview

For experiment 2, males and females were exposed to 0, 100, 200, 300, or 400 Gy x-ray and placed overnight in a 30 ml plastic cup with a non-exposed moth of the opposite sex for an opportunity to mate. Afterward the females was isolated and place in an oviposition jar until death. After death the female was detected to determine if there was a spermatophore, indicating successful mating.

```
library(readxl)
df_mating <- read_excel("Dataset.xlsx", sheet = "Mating_400Gy")
df_mating
```

```
## # A tibble: 92 × 5
##    rep   FemID sex     dose mate
##    <chr> <chr> <chr>  <dbl> <chr>
##  1 c     C1    female     0 No
##  2 c     C2    female     0 Yes
##  3 c     C5    female     0 Yes
##  4 e     E1    female     0 Yes
##  5 e     E2    female     0 Yes
##  6 e     E3    female     0 Yes
##  7 e     E4    female     0 Yes
##  8 e     E5    female     0 Yes
##  9 c     C10   female   100 No
## 10 c     C8    female   100 No
## # ℹ 82 more rows
```

Summary of mating success

Obtain a table indicated number and percent of moths mated.

```
library(dplyr)
tbl_mating <- df_mating %>%
  group_by(sex,dose) %>%
  summarise(nMated = sum(mate == "Yes"),
            nNot = sum(mate == "No"),
            nTotal = sum(!is.na(mate)),
            prop_mated = nMated/nTotal)
write.csv(tbl_mating,"expt2_mating_summary.csv",row.names = F)
tbl_mating
```

```
## # A tibble: 10 × 6
## # Groups:   sex [2]
##    sex     dose nMated  nNot nTotal prop_mated
##    <chr>  <dbl>  <int> <int>  <int>      <dbl>
##  1 female     0      7     1      8      0.875
##  2 female   100      4     3      7      0.571
```

```
##  3 female   200      9     1     10      0.9
##  4 female   300      8     1      9      0.889
##  5 female   400      9     1     10      0.9
##  6 male       0      9     0      9      1
##  7 male     100      7     3     10      0.7
##  8 male     200      5     5     10      0.5
##  9 male     300      7     2      9      0.778
## 10 male     400      9     1     10      0.9
```

Data Analysis

Examine whether there are signficant differenct due to sex or irradiation using a GLM binomial model with both categories categorical.

```
# rename variables for ease of use in formula
tbl_mating2 <- tbl_mating %>%
  rename(y = nMated,
         n = nTotal)

# Make dose a categorical variable
tbl_mating2$dose <- as.factor(tbl_mating2$dose)

# Fit a logistic regression model using glm with binomial family
m1 <- glm(cbind(y,n) ~ sex + dose, data = tbl_mating2, family = binomial)

# Examine model summary for sake of completeness
summary(m1)

##
## Call:
## glm(formula = cbind(y, n) ~ sex + dose, family = binomial, data = tbl_mati
ng2)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.01438    0.38838  -0.037    0.970
## sexmale     -0.08484    0.31502  -0.269    0.788
## dose100     -0.36962    0.52106  -0.709    0.478
## dose200     -0.30503    0.49397  -0.618    0.537
## dose300     -0.12689    0.49400  -0.257    0.797
## dose400     -0.04861    0.47664  -0.102    0.919
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1.63837  on 9  degrees of freedom
## Residual deviance: 0.75263  on 4  degrees of freedom
## AIC: 45.308
##
## Number of Fisher Scoring iterations: 3
```

Using Deviance residual chi-square and df to check model fit

```
pchisq(0.75263, df = 4, lower.tail = F)
```

```
## [1] 0.9446835
```

Get type II Analysis of Deviance

```
library(car)
Anova(m1, type = "II")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: cbind(y, n)
##      LR Chisq Df Pr(>Chisq)
## sex   0.07254  1     0.7877
## dose  0.80817  4     0.9374
```

Experiment 2 – Total Fecundity

Experiment 2 Total Fecundity–Overview

For Experiment 2 eggs were counted at 3 points: soon after they were laid (all eggs, total fecundity), when they turned red (indication of fertility, but complicated by irradiation), and when the headcapsule of a neonate was visible (blackhead stage, killed at this point.). Here we compare total fecundity.

```
library(readxl)
# Load data and display first few rows
df_all <- read_excel("Dataset.xlsx", sheet = "Oviposition_400Gy")

# Set x-ray dose as a factor
df_all$dose <- as.factor(df_all$dose)

library(dplyr)
# Isolate data for females and display the first few rows
females <- df_all %>%
  filter(sex == "female")
females
```

```
## # A tibble: 37 × 5
##    rep   FemID sex    dose  sum_white
##    <chr> <chr> <chr>  <fct>     <dbl>
##  1 c     C2    female 0           225
##  2 c     C5    female 0           241
##  3 e     E1    female 0           135
##  4 e     E2    female 0            61
##  5 e     E3    female 0             8
##  6 e     E4    female 0            41
##  7 e     E5    female 0            81
##  8 c     C7    female 100         263
##  9 c     C9    female 100          75
## 10 e     E10   female 100         147
## # ℹ 27 more rows
```

```
# Isolate data for males and display the first few rows
males <- df_all %>%
  filter(sex == "male")
males
```

```
## # A tibble: 37 × 5
##    rep   FemID sex   dose  sum_white
##    <chr> <chr> <chr> <fct>     <dbl>
##  1 b     B21   male  0           205
##  2 b     B22   male  0           235
##  3 b     B23   male  0           266
##  4 b     B24   male  0           282
##  5 b     B25   male  0           234
```

```
##  6 d     D1    male  0            37
##  7 d     D3    male  0            70
##  8 d     D4    male  0           264
##  9 d     D5    male  0           106
## 10 b     B1    male  100         175
## # ℹ 27 more rows
```

Summary and statistical analysis of female total fecundity

```
library(FSA)            # for SE
# n, mean, and SE by dose for females
females %>%
  group_by(dose) %>%
  summarise(nObs = sum(!is.na(sum_white)),
            mn = mean(sum_white, na.rm = TRUE),
            sem = se(sum_white))
```

```
## # A tibble: 5 × 4
##   dose   nObs    mn   sem
##   <fct> <int> <dbl> <dbl>
## 1 0         7  113.  34.3
## 2 100       4  122   55.4
## 3 200       9  156   26.9
## 4 300       8  150.  33.5
## 5 400       9  126.  36.9
```

We are comparing the sexes separately from the start on the grounds that there is an abundant peer-reviewed literature indicating that difference in radiosensitivity in Lepidoptera is the rule rather than the exception. Starting here with females.

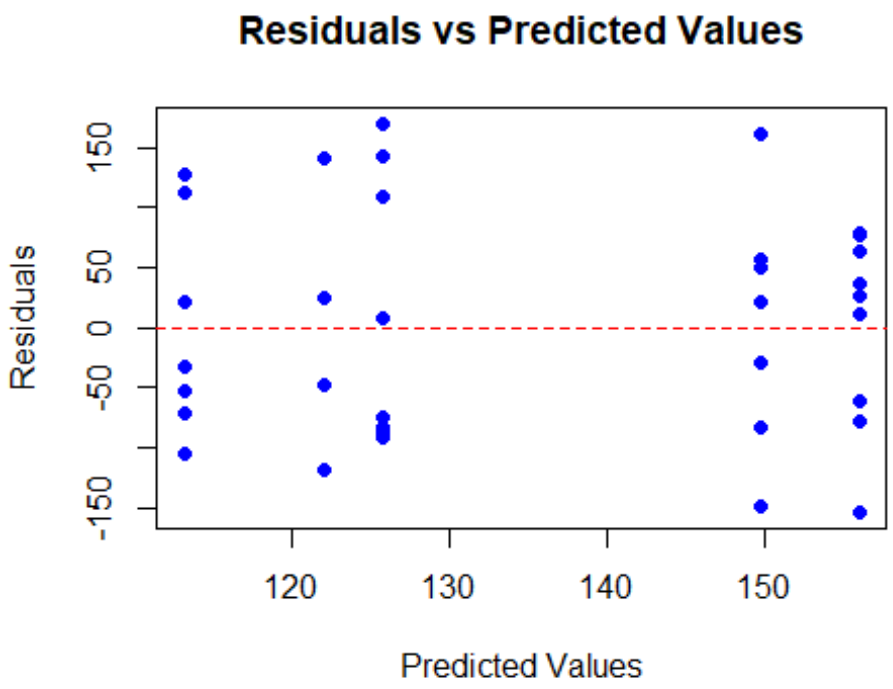
Data distribution: The data above suggest similar means and standard errors between the treatments, which is consistent with traditional ANOVA (Gaussian error distribution). Not shown–the hist() function (native R) and the Desc() function (DescTools) suggest a slight right skew but a box plot of the data that is part of the the Desc() function output suggests that the mean is reasonably centered. Playing with the Poisson, quasi, and quasipoisson family in glm() and negative binomial from MASS finds poor model fit as determined by examining pchisq() for the residual deviance and residual degrees of freedom. The glm() function with a gamma distribution provides good model fit, but does not provide different information compared to the plain ANOVA approach.

```
# m1 -- 1 way ANOVA for females
m1 <- aov(sum_white ~ dose, data = females)
summary(m1)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## dose         4  10485    2621    0.28  0.889
## Residuals   32 299496    9359
```

Now use a residuals plot to check model fit

```r
# Extract the fitted (predicted) values and residuals from the model
pred_values <- fitted(m1)
residuals_val <- residuals(m1)

# Create the residual vs predicted plot
plot(pred_values, residuals_val,
     xlab = "Predicted Values",
     ylab = "Residuals",
     main = "Residuals vs Predicted Values",
     pch = 19, col = "blue")
abline(h = 0, col = "red", lty = 2)
```



Good enough for the present purpose–due diligence done.

Summary and statistical analysis of male total fecundity

```r
# n, mean, and SE by dose for males
males %>%
  group_by(dose) %>%
  summarise(nObs = sum(!is.na(sum_white)),
            mn = mean(sum_white, na.rm = TRUE),
            sem = se(sum_white))
```

```
## # A tibble: 5 × 4
##   dose   nObs    mn   sem
##   <fct> <int> <dbl> <dbl>
## 1 0         9 189.   30.9
```

```
## 2 100       7 189    50.4
## 3 200       5 202.   55.7
## 4 300       7 124.   29.3
## 5 400       9  81.7  26.6
```

After examining several alternatives as described above for the female data set, the MASS package (from the 2002 book “Modern Applied Statistics with S” by Venables and Ripley) was used to apply a GLM with binomial distribution. Model diagnostics indicate that this is the correct model, but no significant differences were found.

```
# Fit GLM with negative binomial
library(MASS)
m2 <- glm.nb(sum_white ~ dose, data = males)
summary(m2)

##
## Call:
## glm.nb(formula = sum_white ~ dose, data = males, init.theta = 1.362871472,
##     link = log)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  5.240571   0.286559  18.288   <2e-16 ***
## dose100      0.001176   0.433235   0.003   0.9978
## dose200      0.066707   0.479433   0.139   0.8893
## dose300     -0.421442   0.433693  -0.972   0.3312
## dose400     -0.837925   0.406206  -2.063   0.0391 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(1.3629) family taken to be 1)
##
##     Null deviance: 48.719  on 36  degrees of freedom
## Residual deviance: 42.807  on 32  degrees of freedom
## AIC: 452
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  1.363
##           Std. Err.:  0.306
##
##  2 x log-likelihood:  -440.001

# Examine fit using Deviance fit
pchisq(42.807,32,lower.tail = F)

## [1] 0.09603374
```

A value of < 0.05 would definitively indicate poor model fit. As it is, negative binomial is a logical choice for over-dispersed count data that includes 0s, and this test suggests model fit is good enough.

Using the Anova() function from the car package to get a type II analysis:

```
library(car)
# Model parameters with type II degrees of freedom
Anova(m2)

## Analysis of Deviance Table (Type II tests)
##
## Response: sum_white
##      LR Chisq Df Pr(>Chisq)
## dose   5.9119  4     0.2058
```

The P value with type II degree of freedom says “not significant”. While the raw model summary indicates P < 0.05 for 400 Gy vs. 0 Gy, that is without the type II adjustment and would of course be lost after adjusting for multiple comparisons.

Experiment 2 – Red Eggs

Experiment 2 Red Eggs–Overview

As mentioned in the previous document, the red egg stage is the second of three developmental stages at which egg viability was assessed. In normal development, there is a near total correlation between egg viability and eggs turning from white to red. Irradiation, however, can cause fatal developmenal problems aftere this color shift which reduced the reliablity of egg color as an indicator of fertility. As previously, the sexes are separated because of expected differences in radiosensitivy and the proportion of red eggs are by x-ray treatment (0, 100, 200, 300, or 400 Gy)

```
library(readxl)
# Load data and display first few rows
df_red <- read_excel("Dataset.xlsx", sheet = "Development_400Gy")

# Set x-ray dose as a factor (categorical)
df_red$dose <- as.factor(df_red$dose)
```

Separate sexes

```
library(dplyr)
# Isolate data for females and display the first few rows
females <- df_red %>%
  filter(sex == "female")
females
```

```
## # A tibble: 37 × 5
##    rep   FemID sex    dose  sum_red
##    <chr> <chr> <chr>  <fct>   <dbl>
##  1 c     C2    female 0         175
##  2 c     C5    female 0         145
##  3 e     E1    female 0         107
##  4 e     E2    female 0          14
##  5 e     E3    female 0           0
##  6 e     E4    female 0           0
##  7 e     E5    female 0           0
##  8 c     C7    female 100       110
##  9 c     C9    female 100         4
## 10 e     E10   female 100        96
## # ℹ 27 more rows
```

```
# Isolate data for males and display the first few rows
males <- df_red %>%
  filter(sex == "male")
males
```

```
## # A tibble: 37 × 5
##    rep   FemID sex   dose  sum_red
```

```
##    <chr> <chr> <chr> <fct>   <dbl>
##  1 b     B21   male  0         198
##  2 b     B22   male  0         208
##  3 b     B23   male  0         180
##  4 b     B24   male  0         240
##  5 b     B25   male  0         217
##  6 d     D1    male  0          25
##  7 d     D3    male  0           4
##  8 d     D4    male  0         179
##  9 d     D5    male  0          19
## 10 b     B1    male  100       160
## # ℹ 27 more rows
```

Summary and statistical analysis of red eggs from irradiated females

```
library(FSA)            # for SE
# n, mean, and SE by dose for females
females %>%
  group_by(dose) %>%
  summarise(nObs = sum(!is.na(sum_red)),
            mn = mean(sum_red, na.rm = TRUE),
            sem = se(sum_red),
            pct_sterile = 100*(1-(sum(sum_red > 0)/nObs)))
```

```
## # A tibble: 5 × 5
##   dose   nObs    mn   sem pct_sterile
##   <fct> <int> <dbl> <dbl>       <dbl>
## 1 0         7 63    29.1         42.9
## 2 100       4 52.5  29.3         25
## 3 200       9 12.3   5.38        11.1
## 4 300       8  7.62  4.56        50
## 5 400       9  2.11  1.14        55.6
```

The mean and SE are consistent with overdispersed count data. The negative binomial is appropriate for such data. The percent sterility calculation reminds us that some females are laying at least some red eggs.

```
# Analysis of Deviance using GLM with negative binomial--females
library(MASS)
m1 <- MASS::glm.nb(sum_red ~ dose, data = females)
summary(m1)
```

```
##
## Call:
## MASS::glm.nb(formula = sum_red ~ dose, data = females, init.theta = 0.2966
718039,
##     link = log)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   4.1431     0.6956   5.957 2.58e-09 ***
```

```
## dose100      -0.1823     1.1538  -0.158 0.874442
## dose200      -1.6308     0.9313  -1.751 0.079926 .
## dose300      -2.1117     0.9600  -2.200 0.027823 *
## dose400      -3.3959     0.9544  -3.558 0.000374 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(0.2967) family taken to be 1)
##
##     Null deviance: 53.093  on 36  degrees of freedom
## Residual deviance: 38.536  on 32  degrees of freedom
## AIC: 246.58
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  0.2967
##           Std. Err.:  0.0784
##
##  2 x log-likelihood:  -234.5770
```

Examine model fit for females

```
# Calculate p-value under the chi-squared distribution
dev <- deviance(m1)
df <- df.residual(m1)
p_value <- pchisq(dev, df, lower.tail = FALSE)
cat("Chi-square goodness-of-fit p-value:", p_value, "\n")
```

```
## Chi-square goodness-of-fit p-value: 0.1978609
```

The residual deviance is not significantly different (P > 0.05) from the expected value for the residual degrees of freedom.

```
library(car)
# Model parameters with type II degrees of freedom
Anova(m1)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: sum_red
##      LR Chisq Df Pr(>Chisq)
## dose   14.557  4   0.005713 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Overall, there were significant differences. Proceeding to comparisons of multiple means.

```
library(emmeans)
library(multcomp)
# Means separation
```

```
emm_fem <- emmeans(m1, ~ dose, type = "response")
multcomp::cld(emm_fem, Letters = LETTERS, decreasing = TRUE)
```

```
##  dose response    SE  df asymp.LCL asymp.UCL .group
##  0       63.00 43.80 Inf    16.117     246.3  A
##  100     52.50 48.30 Inf     8.641     319.0  A
##  200     12.33  7.64 Inf     3.664      41.5  AB
##  300      7.62  5.04 Inf     2.085      27.9  AB
##  400      2.11  1.38 Inf     0.586       7.6   B
##
## Confidence level used: 0.95
## Intervals are back-transformed from the log scale
## P value adjustment: tukey method for comparing a family of 5 estimates
## Tests are performed on the log scale
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
##       then we cannot show them to be different.
##       But we also did not show them to be the same.
```

Summary and statistical analysis of red eggs from females mated with irradiated males

```
# n, mean, and SE by dose for females
males %>%
  group_by(dose) %>%
  summarise(nObs = sum(!is.na(sum_red)),
            mn = mean(sum_red, na.rm = TRUE),
            sem = se(sum_red),
            pct_sterile = 100*(1-(sum(sum_red > 0)/nObs)))
```

```
## # A tibble: 5 × 5
##   dose   nObs    mn   sem pct_sterile
##   <fct> <int> <dbl> <dbl>       <dbl>
## 1 0         9 141.   31.9         0
## 2 100       7 151    47.5        14.3
## 3 200       5 109.   37.6         0
## 4 300       7  39.3  23.4        14.3
## 5 400       9  17.3  12.3        33.3
```

```
# Analysis of Deviance using GLM with negative binomial--males
m2 <- MASS::glm.nb(sum_red ~ dose, data = males)
summary(m2)
```

```
##
## Call:
## MASS::glm.nb(formula = sum_red ~ dose, data = males, init.theta = 0.527194
438,
##     link = log)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
```

```
## (Intercept)  4.94955    0.45994  10.761  < 2e-16 ***
## dose100      0.06773    0.69532   0.097  0.92240
## dose200     -0.25454    0.76990  -0.331  0.74094
## dose300     -1.27869    0.69725  -1.834  0.06667 .
## dose400     -2.09692    0.65476  -3.203  0.00136 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(0.5272) family taken to be 1)
##
##     Null deviance: 56.944  on 36  degrees of freedom
## Residual deviance: 44.726  on 32  degrees of freedom
## AIC: 386.47
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  0.527
##           Std. Err.:  0.118
##
##  2 x log-likelihood:  -374.469
```

Examine model fit for males

```
# Calculate p-value under the chi-squared distribution
dev <- deviance(m2)
df <- df.residual(m2)
p_value <- pchisq(dev, df, lower.tail = FALSE)
cat("Chi-square goodness-of-fit p-value:", p_value, "\n")

## Chi-square goodness-of-fit p-value: 0.06688904
```

The residual deviance is not significantly different (P > 0.05) from the expected value for the residual degrees of freedom.

```
# Model parameters with type II degrees of freedom
Anova(m2)

## Analysis of Deviance Table (Type II tests)
##
## Response: sum_red
##      LR Chisq Df Pr(>Chisq)
## dose   12.218  4     0.0158 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Overall, there were significant differences. Proceeding to comparisons of multiple means.

```
# Model parameters with type II degrees of freedom
emm_males <- emmeans(m2, ~ dose, type = "response")
multcomp::cld(emm_males, Letters = LETTERS, decreasing = TRUE)
```

```
##  dose response    SE  df asymp.LCL asymp.UCL .group
##  100     151.0 78.70 Inf     54.34     419.6  A
##  0       141.1 64.90 Inf     57.29     347.6  A
##  200     109.4 67.50 Inf     32.62     366.9  AB
##  300      39.3 20.60 Inf     14.07     109.7  AB
##  400      17.3  8.08 Inf      6.95      43.2   B
##
## Confidence level used: 0.95
## Intervals are back-transformed from the log scale
## P value adjustment: tukey method for comparing a family of 5 estimates
## Tests are performed on the log scale
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
##       then we cannot show them to be different.
##       But we also did not show them to be the same.
```

Experiment 2 – Blackhead Eggs

Experiment 2 Red Eggs–Overview

Here we examine the third developmental stage in eggs–blackhead (i.e., the headcapsule of the neonate larva is visible). The sexes are again examined separately because of expected differences in radiosensitivy.

```
library(readxl)
# Load data and display first few rows
df_black <- read_excel("Dataset.xlsx", sheet = "Blackhead_400Gy")

# Set x-ray dose as a factor (categorical)
df_black$dose <- as.factor(df_black$dose)
```

Separate sexes

```
library(dplyr)
# Isolate data for females and display the first few rows
females <- df_black %>%
  filter(sex == "female")
females
```

```
## # A tibble: 37 × 5
##    rep   FemID sex    dose  sum_blackhead
##    <chr> <chr> <chr>  <fct>         <dbl>
##  1 c     C2    female 0                 6
##  2 c     C5    female 0                12
##  3 e     E1    female 0                 3
##  4 e     E2    female 0                 0
##  5 e     E3    female 0                 0
##  6 e     E4    female 0                 0
##  7 e     E5    female 0                 0
##  8 c     C7    female 100              23
##  9 c     C9    female 100               0
## 10 e     E10   female 100              17
## # ℹ 27 more rows
```

```
# Isolate data for males and display the first few rows
males <- df_black %>%
  filter(sex == "male")
males
```

```
## # A tibble: 37 × 5
##    rep   FemID sex   dose  sum_blackhead
##    <chr> <chr> <chr> <fct>         <dbl>
##  1 b     B21   male  0                18
##  2 b     B22   male  0                38
##  3 b     B23   male  0                46
##  4 b     B24   male  0                60
```

```
##  5 b     B25   male  0                 9
##  6 d     D1    male  0                 0
##  7 d     D3    male  0                 0
##  8 d     D4    male  0                 7
##  9 d     D5    male  0                 0
## 10 b     B1    male  100              39
## # ℹ 27 more rows
```

Summary and statistical analysis of red eggs from irradiated females

```
library(FSA)            # for SE
# n, mean, and SE by dose for females
females %>%
  group_by(dose) %>%
  summarise(nObs = sum(!is.na(sum_blackhead)),
            mn = mean(sum_blackhead, na.rm = TRUE),
            sem = se(sum_blackhead),
            pct_sterile = 100*(1-(sum(sum_blackhead > 0)/nObs)))
```

```
## # A tibble: 5 × 5
##   dose   nObs     mn   sem pct_sterile
##   <fct> <int>  <dbl> <dbl>       <dbl>
## 1 0         7  3     1.73         57.1
## 2 100       4 10     5.90         50
## 3 200       9  0.333 0.167        66.7
## 4 300       8  0.875 0.515        62.5
## 5 400       9  0.111 0.111        88.9
```

Looks like there are differences, although not necessarily in a logical order.

```
# Analysis of Deviance using GLM with negative binomial--females
library(MASS)
m1 <- MASS::glm.nb(sum_blackhead ~ dose, data = females)
summary(m1)
```

```
##
## Call:
## MASS::glm.nb(formula = sum_blackhead ~ dose, data = females,
##     init.theta = 0.3610803238, link = log)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   1.0986     0.6658   1.650   0.0989 .
## dose100       1.2040     1.0773   1.118   0.2638
## dose200      -2.1972     1.0413  -2.110   0.0349 *
## dose300      -1.2321     0.9656  -1.276   0.2019
## dose400      -3.2958     1.3232  -2.491   0.0127 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(0.3611) family taken to be 1)
```

```
##
##     Null deviance: 46.113  on 36  degrees of freedom
## Residual deviance: 25.870  on 32  degrees of freedom
## AIC: 106.13
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  0.361
##           Std. Err.:  0.166
##
##  2 x log-likelihood:  -94.128
```

Examine model fit for females

```
# Calculate p-value under the chi-squared distribution
dev <- deviance(m1)
df <- df.residual(m1)
p_value <- pchisq(dev, df, lower.tail = FALSE)
cat("Chi-square goodness-of-fit p-value:", p_value, "\n")

## Chi-square goodness-of-fit p-value: 0.7693452
```

No indication of problem with model fit

```
library(car)
# Model parameters with type II degrees of freedom
Anova(m1)

## Analysis of Deviance Table (Type II tests)
##
## Response: sum_blackhead
##      LR Chisq Df Pr(>Chisq)
## dose   20.243  4  0.0004471 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Overall, there were significant differences. Proceeding to comparisons of multiple means.

```
library(emmeans)
library(multcomp)
# Means separation
emm_fem <- emmeans(m1, ~ dose, type = "response")
multcomp::cld(emm_fem, Letters = LETTERS, decreasing = TRUE)

##  dose response    SE  df asymp.LCL asymp.UCL .group
##  100    10.000 8.470 Inf    1.9013     52.60  A
##  0       3.000 2.000 Inf    0.8136     11.06  AB
##  300     0.875 0.612 Inf    0.2222      3.45  AB
##  200     0.333 0.267 Inf    0.0694      1.60   B
##  400     0.111 0.127 Inf    0.0118      1.05   B
##
```

```
## Confidence level used: 0.95
## Intervals are back-transformed from the log scale
## P value adjustment: tukey method for comparing a family of 5 estimates
## Tests are performed on the log scale
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
##       then we cannot show them to be different.
##       But we also did not show them to be the same.
```

Summary and statistical analysis of red eggs from females mated with irradiated males

```
# n, mean, and SE by dose for females
males %>%
  group_by(dose) %>%
  summarise(nObs = sum(!is.na(sum_blackhead)),
            mn = mean(sum_blackhead, na.rm = TRUE),
            sem = se(sum_blackhead),
            pct_sterile = 100*(1-(sum(sum_blackhead > 0)/nObs)))
```

```
## # A tibble: 5 × 5
##   dose   nObs    mn   sem pct_sterile
##   <fct> <int> <dbl> <dbl>       <dbl>
## 1 0         9 19.8   7.54        33.3
## 2 100       7 40.1  19.8         42.9
## 3 200       5 41.2  12.4          0
## 4 300       7  2.71  1.23        28.6
## 5 400       9  1     1           88.9
```

```
# Analysis of Deviance using GLM with negative binomial--males
m2 <- MASS::glm.nb(sum_blackhead ~ dose, data = males)
summary(m2)
```

```
##
## Call:
## MASS::glm.nb(formula = sum_blackhead ~ dose, data = males, init.theta = 0.
3558404849,
##     link = log)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.9846     0.5638   5.294  1.2e-07 ***
## dose100       0.7079     0.8502   0.833 0.405081
## dose200       0.7339     0.9406   0.780 0.435270
## dose300      -1.9860     0.8786  -2.260 0.023796 *
## dose400      -2.9846     0.8609  -3.467 0.000527 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(0.3558) family taken to be 1)
```

```
##
##     Null deviance: 57.840  on 36  degrees of freedom
## Residual deviance: 37.584  on 32  degrees of freedom
## AIC: 238.6
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  0.356
##           Std. Err.:  0.105
##
##  2 x log-likelihood:  -226.596
```

Examine model fit for males

```
# Calculate p-value under the chi-squared distribution
dev <- deviance(m2)
df <- df.residual(m2)
p_value <- pchisq(dev, df, lower.tail = FALSE)
cat("Chi-square goodness-of-fit p-value:", p_value, "\n")

## Chi-square goodness-of-fit p-value: 0.228605
```

The residual deviance is not significantly different (P > 0.05) from the expected value for the residual degrees of freedom.

```
# Model parameters with type II degrees of freedom
Anova(m2)

## Analysis of Deviance Table (Type II tests)
##
## Response: sum_blackhead
##      LR Chisq Df Pr(>Chisq)
## dose   20.255  4  0.0004447 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Overall, there were significant differences. Proceeding to comparisons of multiple means.

```
# Model parameters with type II degrees of freedom
emm_males <- emmeans(m2, ~ dose, type = "response")
multcomp::cld(emm_males, Letters = LETTERS, decreasing = TRUE)

##  dose response     SE  df asymp.LCL asymp.UCL .group
##  200     41.20 31.000 Inf     9.419    180.22  AB
##  100     40.14 25.500 Inf    11.532    139.74  A
##  0       19.78 11.200 Inf     6.550     59.71  AB
##  300      2.71  1.830 Inf     0.725     10.17   BC
##  400      1.00  0.651 Inf     0.279      3.58    C
##
## Confidence level used: 0.95
## Intervals are back-transformed from the log scale
```

```
## P value adjustment: tukey method for comparing a family of 5 estimates
## Tests are performed on the log scale
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
##       then we cannot show them to be different.
##       But we also did not show them to be the same.
```

Experiment 3 – Pupal Development

Experiment 3–Pupae–Overview

A third experiment examined whether emerged neonates observed in Experiment 2 would continue to develop, thereby damaging their host. Males and females irradiated at 250 and 350 Gy were compare to 0 Gy sham treatments. Response variables examined were the proportion of moths that mated, total fecundity, and the number of F1 that developed to the point of cocooning (prepupa or pupa).

Mating

Summarize the proportion of moths mated by dose and sex

```
# Load the data set
library(readxl)
df_mating350 <- read_excel("Dataset.xlsx", sheet = "Mating_350Gy")

# Make dose a categorical variable
df_mating350$Dose <- as.factor(df_mating350$Dose)
df_mating350
```

```
## # A tibble: 30 × 3
##    Dose  Sex   Mated
##    <fct> <chr> <chr>
##  1 0     F     Yes
##  2 0     F     Yes
##  3 0     F     Yes
##  4 0     F     Yes
##  5 0     F     Yes
##  6 250   F     Yes
##  7 250   F     Yes
##  8 250   F     Yes
##  9 250   F     Yes
## 10 250   F     Yes
## # ℹ 20 more rows
```

```
# Summarize data
library(dplyr)
library(tidyr)
tbl_mating3 <- df_mating350 %>%
  group_by(Sex,Dose) %>%
  summarise(nObs = n(),
            nMated = sum(Mated == "Yes"),
            pct_mated = 100*nMated/nObs)
tbl_mating3
```

```
## # A tibble: 6 × 5
## # Groups:   Sex [2]
```

```
##   Sex   Dose   nObs nMated pct_mated
##   <chr> <fct> <int>  <int>     <dbl>
## 1 F     0         5      5       100
## 2 F     250       5      5       100
## 3 F     350       5      5       100
## 4 M     0         5      4        80
## 5 M     250       5      5       100
## 6 M     350       5      4        80
```

Determine if there are significant differences with logistic regression

```
# Apply categorical logistic regression model
m1 <- glm(cbind(nMated,nObs) ~ Sex + Dose, family = binomial, data = tbl_mati
ng3)

# Display basic Analysis of Deviance
summary(m1)

##
## Call:
## glm(formula = cbind(nMated, nObs) ~ Sex + Dose, family = binomial,
##     data = tbl_mating3)
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.644e-02  5.226e-01  -0.070    0.944
## SexM        -1.458e-01  5.266e-01  -0.277    0.782
## Dose250      1.093e-01  6.418e-01   0.170    0.865
## Dose350      3.609e-17  6.502e-01   0.000    1.000
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 0.153703  on 5  degrees of freedom
## Residual deviance: 0.040652  on 2  degrees of freedom
## AIC: 24.642
##
## Number of Fisher Scoring iterations: 3

# Examine model fit using residual deviance and residual degrees of freedom
resid_dev <- deviance(m1)
resid_df <- df.residual(m1)
pchisq(resid_dev, df = resid_df, lower.tail = FALSE)

## [1] 0.9798792
```

Far from 0.05–the model fits

```
library(car)
# Type II test for predictors
Anova(m1)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: cbind(nMated, nObs)
##      LR Chisq Df Pr(>Chisq)
## Sex  0.076718  1     0.7818
## Dose 0.039051  2     0.9807
```

No significant effect on mating due to sex or dose

Effect of 250 and 350 Gy on total fecundity

Total fecundity in this case is actually the total number of eggs laid over 3 days following mating

```
library(dplyr)
library(FSA)                  # for se() function

# Load egg data
df_eggs350 <- read_excel("Dataset.xlsx", sheet = "Oviposition_350Gy")

# Make dose categorical
df_eggs350$dose <- as.factor(df_eggs350$dose)

# Summarize eggs by sex irradiated and x-ray dose
df_eggs350 %>%
  group_by(Sex,dose) %>%
  summarise(nObs = sum(!is.na(Eggs)),
            mn = mean(Eggs, na.rm = TRUE),
            sem = se(Eggs))
```

```
## # A tibble: 6 × 5
## # Groups:   Sex [2]
##   Sex   dose   nObs    mn   sem
##   <chr> <fct> <int> <dbl> <dbl>
## 1 F     0         5 132.   23.2
## 2 F     250       5  91.2  34.8
## 3 F     350       5  39.4  24.7
## 4 M     0         4 144    45.5
## 5 M     250       5  84    35.1
## 6 M     350       4  67.5  35.7
```

Test for significant effects using a GLM with negative binomial distribution

```
# Examine GLM for eggs
m2 <- MASS::glm.nb(Eggs ~ Sex + dose, data = df_eggs350)
summary(m2)
```

```
##
## Call:
## MASS::glm.nb(formula = Eggs ~ Sex + dose, data = df_eggs350,
```

```
##     init.theta = 0.9704783002, link = log)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   4.8452     0.3808  12.724   <2e-16 ***
## SexM          0.1717     0.3875   0.443   0.6578
## dose250      -0.4511     0.4690  -0.962   0.3361
## dose350      -0.9914     0.4816  -2.058   0.0396 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(0.9705) family taken to be 1)
##
##     Null deviance: 36.280  on 27  degrees of freedom
## Residual deviance: 32.117  on 24  degrees of freedom
## AIC: 315.26
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  0.970
##           Std. Err.:  0.237
##
##  2 x log-likelihood:  -305.263
```

```
# Examine model fit
resid_dev2 <- deviance(m2)
resid_df2 <- df.residual(m2)
pchisq(resid_dev2, df = resid_df2, lower.tail = FALSE)
```

```
## [1] 0.1241174
```

Residual deviance not significantly different from expected value

```
# Type II test for female eggs
Anova(m2)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: Eggs
##      LR Chisq Df Pr(>Chisq)
## Sex    0.1955  1     0.6584
## dose   4.0857  2     0.1297
```

Differences in eggs per female between doses are not significant for irradiated females

Pupae per female

Report mean and standard error of F1 cocoons per female by irradiated sex

```r
df_pupae350 <- read_excel("Dataset.xlsx", sheet = "Development_350Gy")

# Rename irradiation variable and make it categorical
names(df_pupae350)[names(df_pupae350) == "TrtLabel"] <- "Dose"
df_pupae350$Dose <- as.factor(df_pupae350$Dose)

# Get cocoons per female by Sex irradiated and dose
df_pupae350 %>%
  group_by(Sex, Dose) %>%
  summarise(nObs = sum(!is.na(Pupae)),
            mn = mean(Pupae, na.rm = TRUE),
            sem = se(Pupae))
```

```
## # A tibble: 6 × 5
## # Groups:   Sex [2]
##   Sex   Dose   nObs    mn    sem
##   <chr> <fct> <int> <dbl>  <dbl>
## 1 F     0         5  48.2  4.91
## 2 F     250       5   0.2  0.2
## 3 F     350       5   2.8  0.970
## 4 M     0         4  60.2 20.2
## 5 M     250       5   1.8  0.490
## 6 M     350       4   2.5  1.19
```

Use GLM with nb to perform Analysis of Deviance examining impact of sex and irradiation

```r
m3 <- MASS::glm.nb(Pupae ~ Sex + Dose, data = df_pupae350)
summary(m3)
```

```
##
## Call:
## MASS::glm.nb(formula = Pupae ~ Sex + Dose, data = df_pupae350,
##     init.theta = 2.241529273, link = log)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   3.8034     0.2713  14.020  < 2e-16 ***
## SexM          0.3914     0.3302   1.185    0.236
## Dose250      -4.0544     0.4503  -9.004  < 2e-16 ***
## Dose350      -2.9888     0.3780  -7.908 2.62e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(2.2415) family taken to be 1)
##
##     Null deviance: 157.257  on 27  degrees of freedom
## Residual deviance:  28.752  on 24  degrees of freedom
## AIC: 161.3
##
## Number of Fisher Scoring iterations: 1
```

```
##
##
##               Theta:  2.242
##           Std. Err.:  0.922
##
##  2 x log-likelihood:  -151.303
```

```
# Examine model fit
resid_dev3 <- deviance(m3)
resid_df3 <- df.residual(m3)
pchisq(resid_dev3, df = resid_df3, lower.tail = FALSE)
```

```
## [1] 0.2296217
```

The ns value indicates no evidence of lack of model fit. Examine type II Analysis of Deviance

```
# Type II test for female eggs
Anova(m3)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: Pupae
##      LR Chisq Df Pr(>Chisq)
## Sex     1.377  1     0.2406
## Dose  128.154  2     <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Examine differences of means for Dose

```
library(emmeans)
library(multcomp)
emm <- emmeans(m3, ~ Dose, type = "response")
multcomp::cld(emm, Letters = LETTERS, decreasing = TRUE)
```

```
##  Dose response     SE  df asymp.LCL asymp.UCL .group
##  0      54.549 12.400 Inf    34.892     85.28  A
##  350     2.746  0.829 Inf     1.520      4.96   B
##  250     0.946  0.366 Inf     0.443      2.02   B
##
## Results are averaged over the levels of: Sex
## Confidence level used: 0.95
## Intervals are back-transformed from the log scale
## P value adjustment: tukey method for comparing a family of 3 estimates
## Tests are performed on the log scale
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
##       then we cannot show them to be different.
##       But we also did not show them to be the same.
```