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.

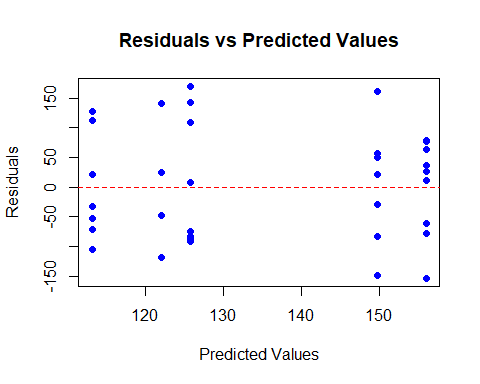
Data distribution: The data above suggest similar means and standard errors between the treatments, which is consistent with traditional ANOVA (Gaussian error distribution). Not shown–the hist() function (native R) and the Desc() function (DescTools) suggest a slight right skew but a box plot of the 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

# 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

# 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.