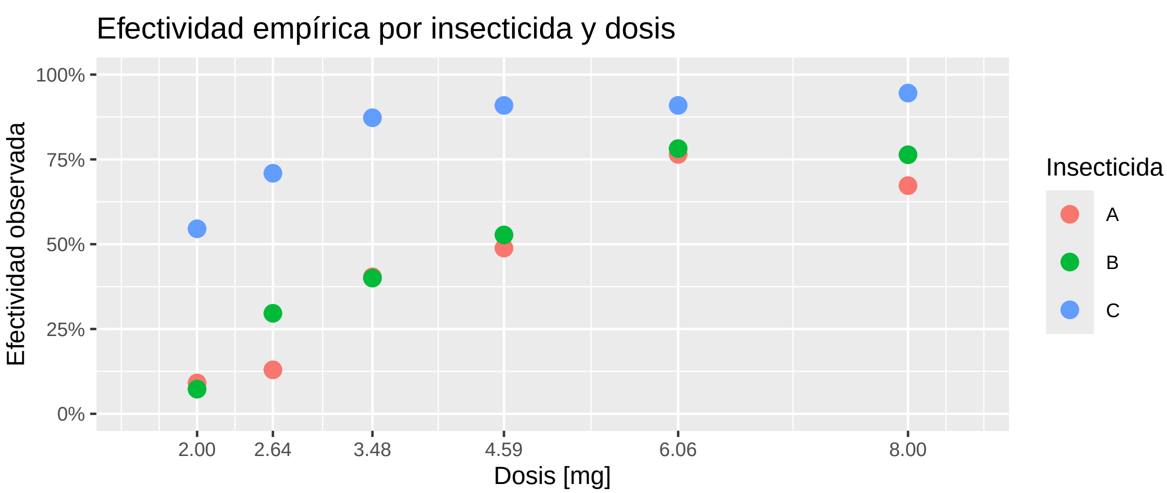
Análisis comparativo de dosis de insecticidas y su impacto en la efectividad

Se analizó la efectividad de los insecticidas A, B y C, probando seis diferentes dosis en diferentes grupos homogeneos de insectos.

Efectividad visualizada en el experimento.

Como primer paso del análisis, se hizo un diagrama de dispersión, que nos ayude a mostrar de manera descriptiva los efectos de cada insecticida por dosis.



Esta primera gráfica es un buen acercamiento para plantear preguntas de investigación, pues para todas las dosis parece que el insecticida C es superior a los insecticidas A, B, los cuales parecen tener un desempeño similar, pues algunos de sus puntos se sobreponen o se encuentran con efectividades muy parecidas.

Modelando con interacciones simples.

Como primer acercamiento, se ajustaron modelos lineales generalizados incluyendo las covariables $Insecticide$, $Deposit$ y sus interacciones.

El componente lineal de dichos modelos se ve como sigue:

$$\eta(Insecticide, Deposit) = \beta_0 + \beta_1 \cdot Deposit + \beta_2 \cdot \mathbb{I}_B + \beta_3 \cdot \mathbb{I}_C + \beta_4 \cdot Deposit \cdot \mathbb{I}_B + \beta_5 \cdot Deposit \cdot \mathbb{I}_C$$

Y para cada insecticida se ve como:

$$\begin{aligned}
\eta(Insecticide = A, Deposit) &= \beta_0 + \beta_1 \cdot Deposit \\
\eta(Insecticide = B, Deposit) &= (\beta_0 + \beta_2) + (\beta_1 + \beta_4) \cdot Deposit \\
\eta(Insecticide = C, Deposit) &= (\beta_0 + \beta_3) + (\beta_1 + \beta_5) \cdot Deposit
\end{aligned}$$

Se ajustaron diferentes modelos considerando el componente lineal anterior, variando las funciones liga (*logit*, *probit* y *c-log-log*), y los resultados se muestran en la tabla siguiente:

Se realizó verificación de supuestos para el modelo de menor AIC, (liga logit) y concluimos que es plausible su uso, pero es mejor considerar más opciones antes de contestar preguntas.

Cuadro 1: Comparación de AIC para Modelos con Interacción Lineal.

GLM	AIC
Bernoulli, liga Logit	999.19
Bernoulli, liga Probit	1000.78
Bernoulli, liga C-log-log	1013.28

Modelando interacciones más complejas.

Si ahora consideramos modelos incluyendo la variable $Deposit^2$, el componente lineal de nuestro GLM se ve de la forma:

$$\begin{aligned}\eta(Insecticide, Deposit) =& \beta_0 + \beta_1 \cdot Deposit + \beta_2 \cdot Deposit^2 + \beta_3 \cdot \mathbb{I}_B + \beta_4 \cdot \mathbb{I}_C + \beta_5 \cdot Deposit \cdot \mathbb{I}_B + \beta_6 \cdot Deposit \cdot \mathbb{I}_C + \\ &+ \beta_7 \cdot Deposit^2 \cdot \mathbb{I}_B + \beta_8 \cdot Deposit^2 \cdot \mathbb{I}_C\end{aligned}$$

Y para cada insecticida obtenemos que:

$$\begin{aligned}\eta(Insecticide = A, Deposit) &= \beta_0 + \beta_1 \cdot Deposit + \beta_2 \cdot Deposit^2 \\ \eta(Insecticide = B, Deposit) &= (\beta_0 + \beta_3) + (\beta_1 + \beta_5) \cdot Deposit + (\beta_2 + \beta_7) \cdot Deposit^2 \\ \eta(Insecticide = C, Deposit) &= (\beta_0 + \beta_4) + (\beta_1 + \beta_6) \cdot Deposit + (\beta_2 + \beta_8) \cdot Deposit^2\end{aligned}$$

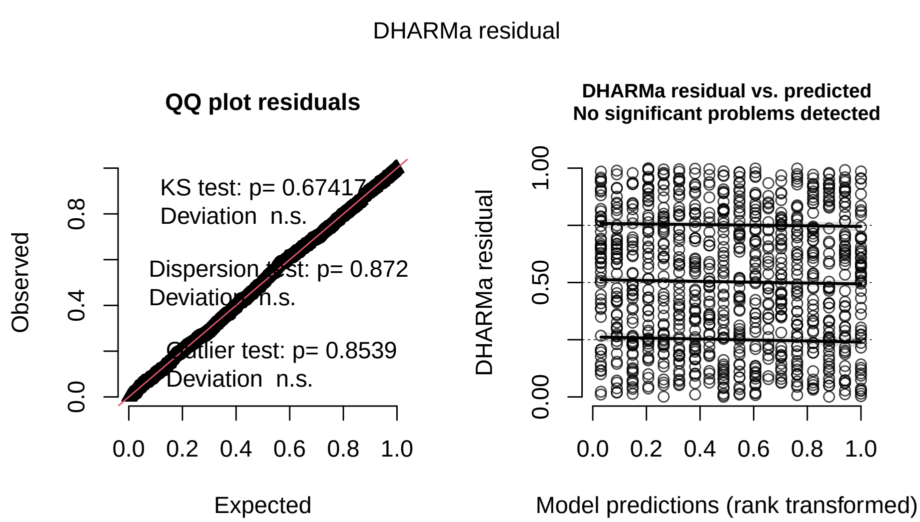
Ajustando modelos con ese componente lineal para las mismas ligas (*logit*, *probit* y *c-log-log*) obtuvimos los resultados de la tabla siguiente:

Cuadro 2: Comparación de AIC para Modelos con Interacción Cuadrática

GLM	AIC
Bernoulli, liga Logit	975.57
Bernoulli, liga Probit	975.72
Bernoulli, liga C-log-log	976.98

Notamos que todos estos modelos son más verosimiles que los de interacciones lineales, y nuevamente, el modelo con la liga logit es el de menor AIC y es nuestro mejor candidato para modelar. Analizando los supuestos de este modelo:

Para nuestro modelo, se rechazo la hipótesis nula en la prueba F asociada a la tabla ANOVA, indicando que este modelo nos aporta información para modelar a la esperanza. Analizando los gráficos de bondad de ajuste y del componente lineal:



- Los puntos se ven muy uniformes sobre la recta del QQ-Plot y ambos tests (de bondad y modelado de varianza) se rechazan, indicando que no hay problemas con estas cuestiones.

- La gráfica del componente lineal se ve muy uniforme, y el test asociado no es significativo, indicando que no tenemos problemas serios con esta parte.

Finalmente, efectuamos una prueba de modelos anidados, con fin de comparar nuestro modelo de interacciones cuadráticas con el de interacciones simples y determinar si es necesario usar el modelo completo.

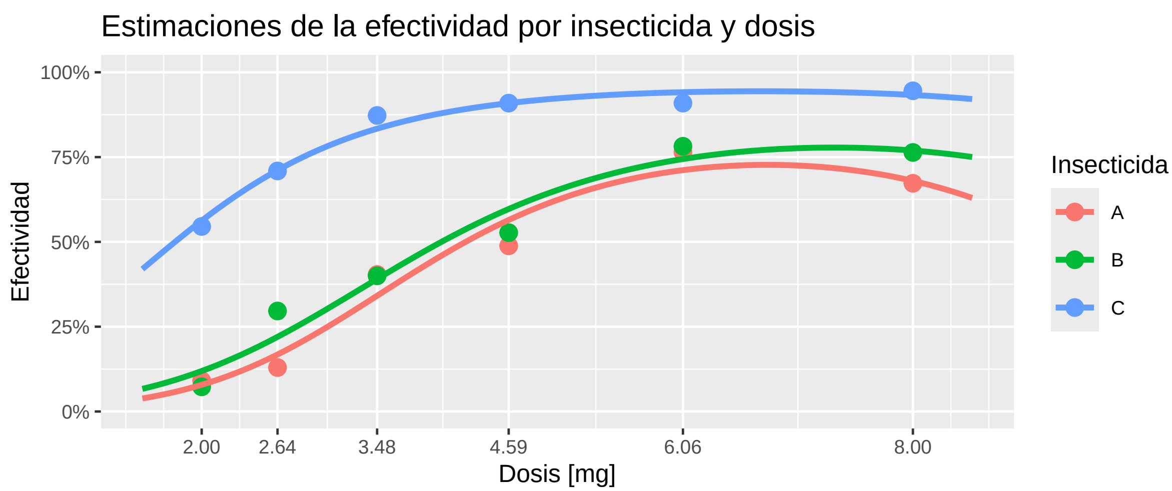
$$H_0 : \text{"Es plausible el modelo reducido"} \qquad vs \qquad H_a : \text{"Se requiere el modelo completo"}$$

De dicha prueba obtuvimos p-value = 1.662e-06 < 0.05, y por tanto, rechazamos H_0. Concluimos que debemos ir por el modelo completo.

Dado que el modelo completo pasó las pruebas de hipótesis, se probó que es significativamente mejor para modelar que nuestro primer modelo candidato y a pesar de tener más parámetros, resultó ser el modelo de menor AIC (más verosimil), será con el que modelemos. GLM Bernoulli, liga Logit, con interacciones cuadráticas.

Estimaciones, comparaciones y conclusión.

A continuación, presentamos nuevamente el gráfico de dispersión del inicio, pero esta vez incluyendo las estimaciones puntuales hechas por nuestro mejor modelo.



Un insecticida efectivo con dosis baja y significativamente mejor.

Lo más importante para un insecticida es que tenga una alta efectividad, así que, calculamos la dosis mínima para matar al 72% de los insectos y se presenta en la tabla siguiente:

Cuadro 3: Dosis Mínima para un 72% de Mortalidad por Insecticida.

Insecticida	Dosis Mínima
A	6.29mg
B	5.69mg
C	2.69mg

Notamos que, la dosis del insecticida C de 2.69mg es mucho menor que las de los otros insecticidas para conseguir ese grado de efectividad(para A y B es más del doble de ese valor).

Considerando el valor de dicha dosis, 2.69mg, podemos hacer una prueba de hipótesis y concluir estadísticamente si para dicho valor, existe un insecticida que es mejor. Como principal candidato tenemos al insecticida C, así que nos interesa verificar el siguiente par de hipótesis:

$$\mathbb{P}(y = 1; Insecticide = C, Deposit = 2.69) > \mathbb{P}(y = 1; Insecticide = A, Deposit = 2.69)$$
$$\mathbb{P}(y = 1; Insecticide = C, Deposit = 2.69) > \mathbb{P}(y = 1; Insecticide = B, Deposit = 2.69)$$

Aplicando la función liga, podemos expresar la prueba de hipótesis global en términos del componente lineal y se ven como sigue:

$$H_0 : [\, \beta_4 + \beta_6 \cdot 2.69 + \beta_8 \cdot 2.69^2 \leq 0 \,] \wedge [\, (\beta_4 - \beta_3) + (\beta_6 - \beta_5) \cdot 2.69 + (\beta_8 - \beta_7) \cdot 2.69^2 \leq 0 \,]$$

$$vs$$

$$H_a : [\, (\beta_4 + \beta_6 \cdot 2.69 + \beta_8 \cdot 2.69^2 > 0 \,] \vee [\, (\beta_4 - \beta_3) + (\beta_6 - \beta_5) \cdot 2.69 + (\beta_8 - \beta_7) \cdot 2.69^2 > 0 \,]$$

Efectuando pruebas de hipótesis simultaneas, obtuvimos que se rechazaron ambas hipótesis nulas ($\leq$) con p-value ajustado menor que 0.05. Podemos garantizar que el insecticida C es mejor que los insecticidas A y B.

Comportamiento de los insecticidas A y B.

Visto lo anterior, ahora podemos hacer una prueba de hipótesis para determinar si los insecticidas A y B tienen comportamientos similares. Planteando las hipótesis en términos de los β's se reducen a la prueba lineal general siguiente:

$$H_0 : (\beta_3 = 0) \wedge (\beta_5 = 0) \wedge (\beta_7 = 0) \qquad vs \qquad H_a : (\beta_3 \neq 0) \wedge (\beta_5 \neq 0) \wedge (\beta_7 \neq 0)$$

El resultado de la prueba fue un p-value $= 0.462$, por tanto, no podemos rechazar H_0 y no hay evidencia para garantizar que los insecticidas A y B se comportan diferente. Por falta de evidencia, podemos asumir que tienen un rendimiento similar.