Tarea 3

Análisis de datos categóricos — Semestre 2020-1

Vega Martínez, Raymundo[1]; Sulca Cavero, Pavel David[1]; Naveja Romero José de Jesús[1]; Dorantes Aldama, Alejandro[1]; y Barba Galdámez, David Francisco[1]

[1]Especialización en Estadística Aplicada, UNAM

19 de marzo de 2020

1. En Estados Unidos se levantó una encuesta para evaluar si la religiosidad de las personas estaba asociada con su preferencia política. Se eligieron 350 personas al azar. Entre las preguntas que se les realizaron estuvieron: ¿Cree usted en el infierno? y ¿Con qué preferencia política se identifica? Los resultados se han resumido en la Tabla 1.

	Sí creo	**No creo**	**No sé**	
Republicano	95	33	10	138
Sin partido	30	20	6	56
Demócrata	80	66	10	156
	205	119	26	350

Tabla 1: Tabla de contingencia de preferencia política vs. religiosidad

a) ¿A qué tipo de esquema de muestreo pertenece este problema? Justifiquen su respuesta.
R: Multinomial, porque se fijó el número total de observaciones ($n = 350$).

b) ¿Es un modelo de independencia o de homogeneidad? Justifiquen su respuesta.
R: Considerando el esquema de muestreo (multinomial), el modelo es de independencia. En otras palabras, evalúa que la probabilidad conjunta es el producto de las probabilidades marginales.

c) Establezcan las hipótesis de la prueba de χ^2.
R: Considerando el contexto de la prueba, las hipótesis de la prueba χ^2 de independencia serían las siguientes:

1) H_0**: La preferencia política de los estadounidenses es independiente de sus creencias religiosas sobre el infierno.**

2) H_a**: La preferencia política de los estadounidenses está asociada a sus creencias religiosas sobre el infierno.**

d) Realicen "a mano" la prueba de χ^2 con y sin corrección de Yates. Concluyan.
R: Con base en la información presentada en la Tabla 1 obtuvimos los siguientes estimadores:

$$\hat{p}_{1\cdot} = \frac{138}{350} = 0.394;\ \hat{p}_{2\cdot} = \frac{56}{350} = 0.16;\ \hat{p}_{3\cdot} = \frac{156}{350} = 0.446$$

$$\hat{p}_{\cdot 1} = \frac{205}{350} = 0.586;\ \hat{p}_{\cdot 2} = \frac{119}{350} = 0.34;\ \hat{p}_{\cdot 3} = \frac{26}{350} = 0.074$$

Con esta información calculamos la frecuencia esperada por independencia (ver Tabla 2).

Así, se puede calcular

$$\chi^2 = \sum_{i,j} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} = \frac{(95 - 80.829)^2}{80.829} + \frac{(30 - 32.8)^2}{32.8} + \frac{(80 - 91.371)^2}{91.371} + \frac{(33 - 46.92)^2}{46.92} + \frac{(20 - 19.04)^2}{19.04}$$
$$+ \frac{(66 - 53.04)^2}{53.04} + \frac{(10 - 10.251)^2}{10.251} + \frac{(6 - 4.16)^2}{4.16} + \frac{(10 - 11.589)^2}{11.589} = 12.52128$$

	Sí creo	No creo	No sé
Republicano	80.829	46.92	10.251
Sin partido	32.8	19.04	4.16
Demócrata	91.371	53.04	11.589

Tabla 2: Frecuencias esperadas por independencia en la Tabla 1.

Utilizando corrección por continuidad de Yates

$$\chi^2_Y = \sum_{i,j} \frac{(|o_{ij} - e_{ij}| - .5)^2}{e_{ij}} = \frac{(|95 - 80.829| - .5)^2}{80.829} + \frac{(|30 - 32.8| - .5)^2}{32.8} + \frac{(|80 - 91.371| - .5)^2}{91.371}$$
$$+ \frac{(|33 - 46.92| - .5)^2}{46.92} + \frac{(|20 - 19.04| - .5)^2}{19.04} + \frac{(|66 - 53.04| - .5)^2}{53.04}$$
$$+ \frac{(|10 - 10.251| - .5)^2}{10.251} + \frac{(|6 - 4.16| - .5)^2}{4.16} + \frac{(|10 - 11.589| - .5)^2}{11.589} = 11.08348$$

Se sabe que este estadístico tiene una distribución ji-cuadrada con $(I-1)(J-1)$ grados de libertad, es decir que, para nuestro caso de estudio se tiene $\chi^2_(4)$. Así, se observa que los valores calculados con y sin corrección de Yates son "grandes", lo que nos llevaría a rechazar la hipótesis nula de que la preferencia política de los estadounidenses es independiente de sus creencias religiosas sobre el infierno. Otra opción es calcular el *p-value* con el siguiente código de R:

```
1-pchisq(12.52128,4) #sin corrección de Yates
1-pchisq(11.08348,4) #con corrección de Yates
```

Los *p-value* calculados son 0.0139 y 0.0256 para las pruebas sin y con corrección de Yates, respectivamente. Si se considera un nivel de significancia de $\alpha = 0.05$, en ambos casos se rechaza la hipótesis nula de independencia.

e) Corroboren el resultado en R. Muestren el resultado.
R: Podemos utilizar el siguiente código para corroborar el resultado en R:

```
CREE <- c("SI", "SI","SI","NO", "NO","NO","NO SE","NO SE","NO SE")
PARTIDO <- c("REP", "SIN","DEM","REP","SIN","DEM","REP","SIN","DEM")
conteos <- c(95,30,80,33,20,66,10,6,10)
TABLA <- data.frame(PARTIDO, CREE, conteos)
DAT<-xtabs(conteos ~ PARTIDO+CREE, data = TABLA)
DAT
chisq.test(DAT,correct = F) #sin corrección de Yates
chisq.test(DAT,correct = T)  #con corrección de Yates
```

Los resultados sin corrección de Yates son los siguientes:

```
> DAT
       CREE
PARTIDO NO NO SE SI
    DEM 66    10 80
    REP 33    10 95
    SIN 20     6 30
> chisq.test(DAT,correct = F)

        Pearson's Chi-squared test

data:  DAT
X-squared = 12.521, df = 4, p-value = 0.01387
```

Se observa que los resultado son idénticos a los obtenidos en la prueba hecha "a mano". Por otra parte, los resultados obtenidos con la corrección de Yates fueron:

```
Pearson's Chi-squared test
data:  DAT
X-squared = 12.521, df = 4, p-value = 0.01387

Warning message:
In chisq.test(DAT, correct = T) :
  Chi-squared approximation may be incorrect
```

En este caso se obtienen los mismos resultados que sin la corrección de Yates. Esto se debe a que, según la *ayuda* de R, la corrección de Yates sólo se aplica a tablas de contingencia de 2x2.

f) Realicen el *mosaicplot* correspondiente. Descríbanlo.

R: El gráfico de mosaicos que se muestra en la Figura 1 indica que casi todas las celdas de la tabla de frecuencias observados (Tabla 1) contienen un valor cercano al esperado por independencia (Tabla 2), solamente los apartidistas tienen una frecuencia de religiosidad un poco menor a la esperada por independencia.

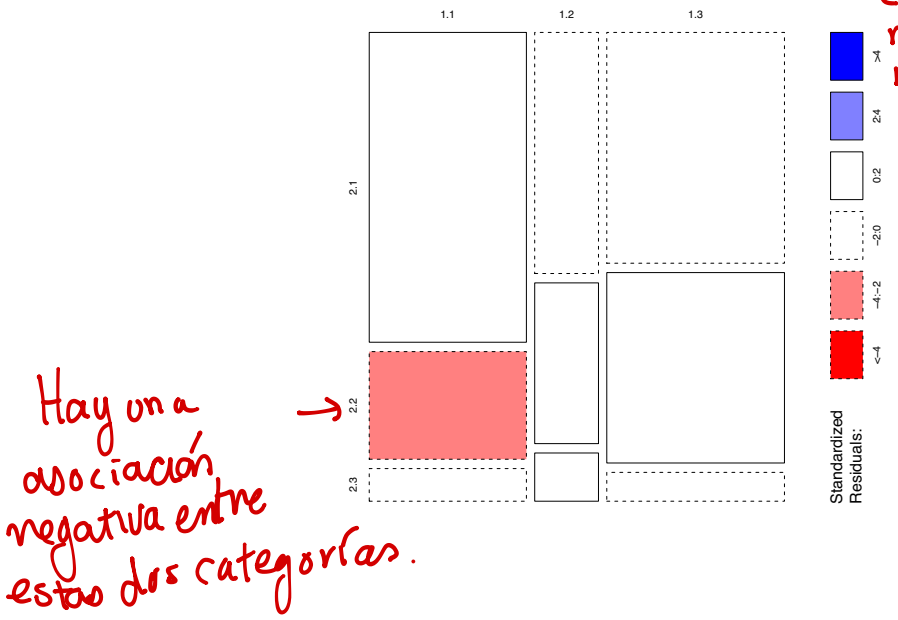

Figura 1: Gráfico de mosaicos para los datos en la Tabla 1.

2. Lean el artículo de Tapia, José A. y Nieto, F. Javier. "Razón de posibilidades: una propuesta de traducción de la expresión *odds ratio*". Salud Pública de México. Julio-Agosto, 1993. Vol. 35 No.4 Pág. 419-424. Elaboren un resumen en media cuartilla.

R:

La expresión inglesa *odds ratio* es ampliamente utilizada en bioestadística y epidemiología. Sin embargo, no existe aún consenso para su traducción. Entre las propuestas planteadas por diversos autores se encuentran "razón de probabilidad", "desigualdad relativa" y "razón de momios". Dicha ambigüedad en la traducciones se debe al poco rigor con que se ha definido el término *odds ratio* en algunos documentos de salud pública y de estadística. En inglés, la palabra *odds* es un sustantivo que hace referencia al cociente entre la probabilidad de que algo ocurra (P) y la probabilidad de que no ocurra ($1 - P$), es decir, al cociente entre el número de posibles "éxitos" y el de posibles "fracaso", si todos los resultados son igualmente posibles. De acuerdo a su definición, no existe una cota superior para el valor de las *odds*, pudiendo ser mayores a 1. En consecuencia, la expresión *odds ratio* hace referencia a la razón entre las *odds* a favor de que un evento ocurra en determinadas circunstancias o que suceda en otras.

Con base en lo anterior, la expresión "razón de probabilidad" es incorrecta debido a que los *odds* no son probabilidades; mientras que "desigualdad relativa", pese a ser intuitivamente correcta, también alude a la desigualdad relativa de los dos grupos que se comparan por lo que no debería utilizarse. Por otra parte, la expresión "razón de momios", pese a ser correcta, posee el inconveniente de utilizar el localismo mexicano *momio* que puede generar confusión en su uso fuera del país.

La traducción alternativa para *odds ratio* propuesta por los autores (José y Javier, 1993) es "razón de posibilidades". De acuerdo a ellos, el término "posibilidades" se usa en español con un sentido bastante similar al significado matemático de *odds*. No obstante, los autores concuerdan con Porta Serra (1990) de que cuando se traduzca la expresión *odds ratio* en textos científicos se debe mencionar entre paréntesis en inglés para evitar posibles confusiones.

3. En el artículo Clopper, C. J. & Pearson, E. S. 1934. The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial Biometrika, Vol. 26, No. 4, vienen las tablas para calcular un intervalo exacto de confianza para p. Utilicen esas tablas para calcular un intervalo al 95 % y otro al 99 % para una muestra tamaño $n = 50$ y $x = 9$. Muestren las tablas y señalen los valores calculados en ellas.

R: Para una muestra de tamaño $n = 50$ y $x = 9$, se tiene:

$$\hat{p} = \frac{9}{50} = 0.18$$

Por lo que evaluando las tablas propuestas por (Clopper y Pearson, 1934) se obtuvieron los siguientes intervalos de confianza (Fig. 2 y 3):

***a*) Intervalo del 95 % confianza:** $p \in (0.085\ ,\ 0.317)$**.**

***b*) Intervalo del 99 % confianza:** $p \in (0.063\ ,\ 0.357)$**.**

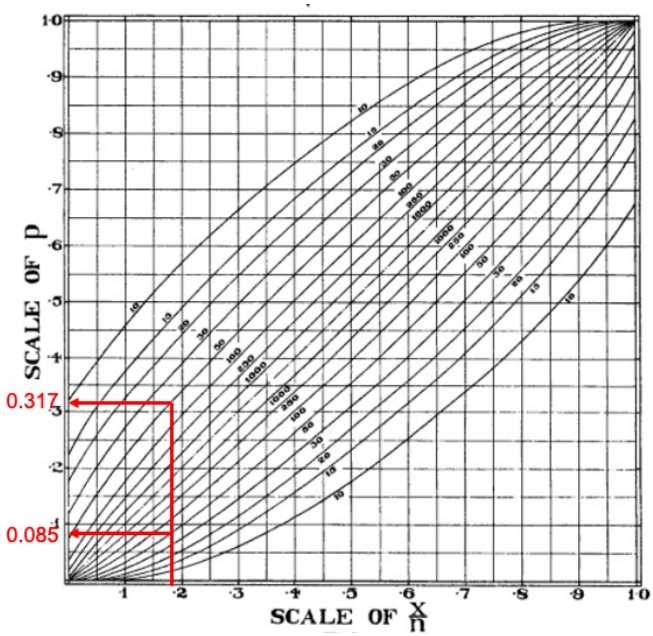

Figura 2: Determinación de intervalo de confianza del 95 %.

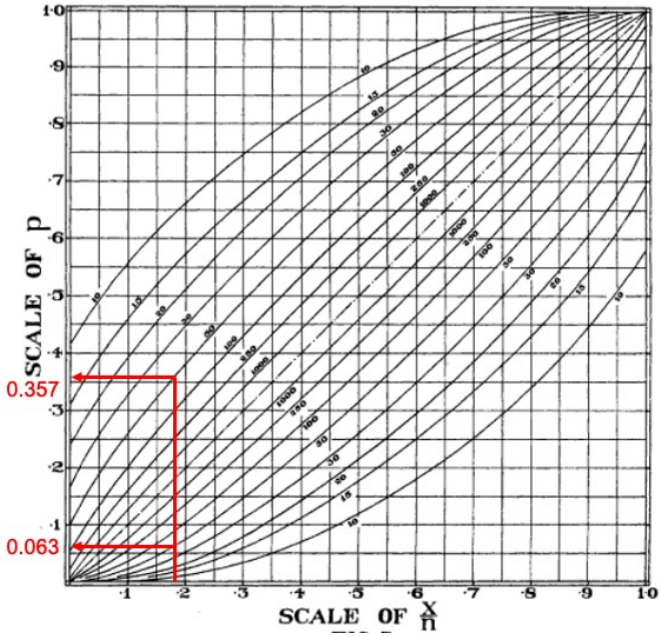

Figura 3: Determinación de intervalo de confianza del 99 %.

4. En una escuela primaria se le pregunta a los niños si les gusta leer. Se seleccionan 15 niños a los que sí les gusta leer y 15 niños a los que no les gusta leer. A su vez, se les pregunta si les gustan las matemáticas. Los resultados se encuentran resumidos en la Tabla 4.

	Sí le gusta leer	No le gusta leer
Sí le gustan las matemáticas	12	2
No le gustan las matemáticas	3	13

a) ¿A qué tipo de esquema de muestreo pertenece este problema? Justifiquen su respuesta.
R: Multinomial-producto, porque se han fijado el total de marginales por columna ($n_{\cdot i} = 15$).

b) ¿Es un modelo de independencia o de homogeneidad? Justifiquen su respuesta.
R: Considerando el esquema de muestreo, es un modelo de homogeneidad.

c) Realiza la prueba exacta de Fisher para estos datos. Interprétala.
R: Efecutamos la prueba exacta de Fisher en R a partir del siguiente código:

```
(tabla1 <-
    matrix(c(12,3,2,13),byrow = TRUE,
           nrow = 2,
           dimnames = list( matematicas =c("SI", "NO"),leer = c("SI", "NO"))))
fisher.test(tabla1)
```

Los resultados obtenidos fueron:

```
        Fisher's Exact Test for Count Data
data:  tabla1
p-value = 0.0006789
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
   2.896661 310.333793
sample estimates:
odds ratio
  22.12942
```

Así, se obtiene un valor $p < .001$. Por tanto, hay evidencia para rechazar la hipótesis nula y concluir que el gusto por las matemáticas varía en niños con o sin gusto por la lectura.

5. Los datos contenidos en el archivo semillas_bosque.csv corresponden a un experimento que llevaron a cabo unos biólogos. Sobre un camino en el bosque (0 metros), colocaron 50 semillas de cacahuate, 50 semillas de girasol y 50 semillas de frijol. Hicieron lo mismo a 15 y 30 metros del camino, como se esquematiza a continuación. Al día siguiente, contaron el número de semillas que habían quedado (semillas.presentes) y el número de semillas que se habían llevado los animales del bosque (semillas.removidas).

a) Importa los datos a R y especifica los siguientes modelos:

```
datos <- read.csv("archivo.csv")

modelo1 <- glm(cbind(semillas.removidas,semillas.presentes)~especie,
               binomial(link="logit"), data=datos)

modelo2 <- glm(cbind(semillas.removidas,semillas.presentes)~ distancia.camino.m,
               binomial(link="logit"), data=datos)

modelo3 <- glm(cbind(semillas.removidas,semillas.presentes)~especie+distancia.camino.m,
               binomial(link="logit"),data=datos)
```

p = probabilidad de que la semilla sea removida.

R: Definimos la función $logit : \mathbb{R}^+ \to \mathbb{R}$ como $logit(p) \equiv log(p/(1-p))$. Así, los resultados obtenidos con el modelo 1 fueron los siguientes:

```
> summary(modelo1)

Call:
glm(formula = cbind(semillas.removidas, semillas.presentes) ~
    especie, family = binomial(link = "logit"), data = datos)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
```

```
-9.898  -5.594  -0.765   6.861   9.529

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)     0.50905    0.04215   12.08   <2e-16 ***
especiefrijol  -0.74345    0.05888  -12.63   <2e-16 ***
especiegirasol -0.90064    0.05923  -15.21   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 5507.6  on 143  degrees of freedom
Residual deviance: 5237.0  on 141  degrees of freedom
AIC: 5603.1

Number of Fisher Scoring iterations: 4
```

Para el modelo 2 obtuvimos:

```
> summary(modelo2)

Call:
glm(formula = cbind(semillas.removidas, semillas.presentes) ~
    distancia.camino.m, family = binomial(link = "logit"), data = datos)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-9.6686  -5.3747  -0.7404   6.6361   9.9338

Coefficients:
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)        -0.52038    0.03835  -13.57   <2e-16 ***
distancia.camino.m  0.03188    0.00198   16.10   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 5507.6  on 143  degrees of freedom
Residual deviance: 5240.9  on 142  degrees of freedom
AIC: 5605

Number of Fisher Scoring iterations: 4
```

Y para el modelo 3:

```
> summary(modelo3)

Call:
glm(formula = cbind(semillas.removidas, semillas.presentes) ~
    especie + distancia.camino.m, family = binomial(link = "logit"),
    data = datos)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-11.0222   -4.5815   -0.8067    5.5311   10.6369

Coefficients:
                    Estimate Std. Error z value Pr(>|z|)
(Intercept)         0.031668   0.051210   0.618    0.536
especiefrijol      -0.773649   0.060146 -12.863   <2e-16 ***
especiegirasol     -0.937169   0.060543 -15.479   <2e-16 ***
distancia.camino.m  0.033196   0.002025  16.392   <2e-16 ***
```

```
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 5507.6  on 143  degrees of freedom
Residual deviance: 4959.6  on 140  degrees of freedom
AIC: 5327.8

Number of Fisher Scoring iterations: 4
```

b) Escribe la fórmula completa de cada uno de estos modelos.

R:

Modelo 1: $logit(p) = 0.5090 - 0.7434 \times \text{frijol} - 0.9006 \times \text{girasol}$

Modelo 2: $logit(p) = -0.52038 + 0.03189 \times \text{distancia (m)}$

Modelo 3: $logit(p) = 0.03167 - 0.77365 \times \text{frijol} - 0.93717 \times \text{girasol} + 0.03320 \times \text{distancia (m)}$

c) Ejecuta la instrucción anova(modelo3, test=Çhisq"). Explícala, indicando en un esquema el valor de χ^2 crítico, devianza, *p-value* y α para cada una de las comparaciones de la tabla. (Nota: Los esquemas pueden ser realizados a mano, en R o de otra manera. Sólo interesa que sean claros y legibles, y que se note que han entendido la utilidad de los valores en la tabla de análisis de devianza.)

R: Como se muestra en la Tabla 3, el modelo especie+distancia explica mucho mejor los datos que el modelo especie o el modelo nulo. Lo anterior puede observarse en la reducción de la devianza de la segunda columna. Sin embargo, si utilizamos a la devianza como estadístico de bondad de ajuste (fijando el valor de la significancia en todos los casos como $\alpha = 0.05$.), vemos que ninguno de los modelos proporciona un ajuste adecuado a los datos.

Variable	Devianza	gl	valor p	χ^2_{crit}	Dif. devianzas	gl	valor p	χ^2_{crit}
nulo	5507.6	143	≈ 0	171.9	—	—	—	—
especie	5237.0	141	≈ 0	169.7	270.6	2	$\approx 1.7 \times 10^{-59}$	5.99
distancia (m)	4959.6	140	≈ 0	168.6	277.4	1	$\approx 2.8 \times 10^{-62}$	3.84

Tabla 3: Análisis de varianza del modelo 3

En la Figura 4, se esquematizan los resultados de la tabla anterior, en el contexto de la distribución χ^2 que les corresponde.

d) ¿Cuál es el mejor modelo y por qué? Utiliza la devianza, el AIC, la significancia de los coeficientes y el análisis de devianza que ejecutaste en el punto 3 para justificar tu respuesta.

R: En la Tabla 4 se muestran algunos estadísticos de bondad de ajuste, que indican que el modelo 3 es mejor. Además, los valores p de las variables en los tres modelos son muy bajos $<< .001$. Por lo tanto, el modelo 3 parece ser el más apropiado, porque ambas variables aportan información independiente. Pero, como se mencionó en el inciso anterior, ninguno de los tres modelos tiene un ajuste adecuado si consideramos el criterio de la devianza.

Modelo	Devianza	AIC	Dif. devianzas
1	5237	5603	270.64
2	5241	5605	266.73
3	4960	5328	270.64, 277.38

Tabla 4: Comparación de estadísticos de bondad de ajuste de los tres modelos

e) ¿Qué medida conviene calcular para evaluar la bondad de ajuste en datos agrupados y cuál conviene para datos desagrupados?

R: Para datos agrupados, se puede utilizar la devianza directamente. En el caso de datos desagrupados, se puede utilizar la estadística de Hosmer-Lemeshow.

f) ¿Estos son datos agrupados o desagrupados? Evalúa la bondad de ajuste en el modelo que elegiste. Explica.

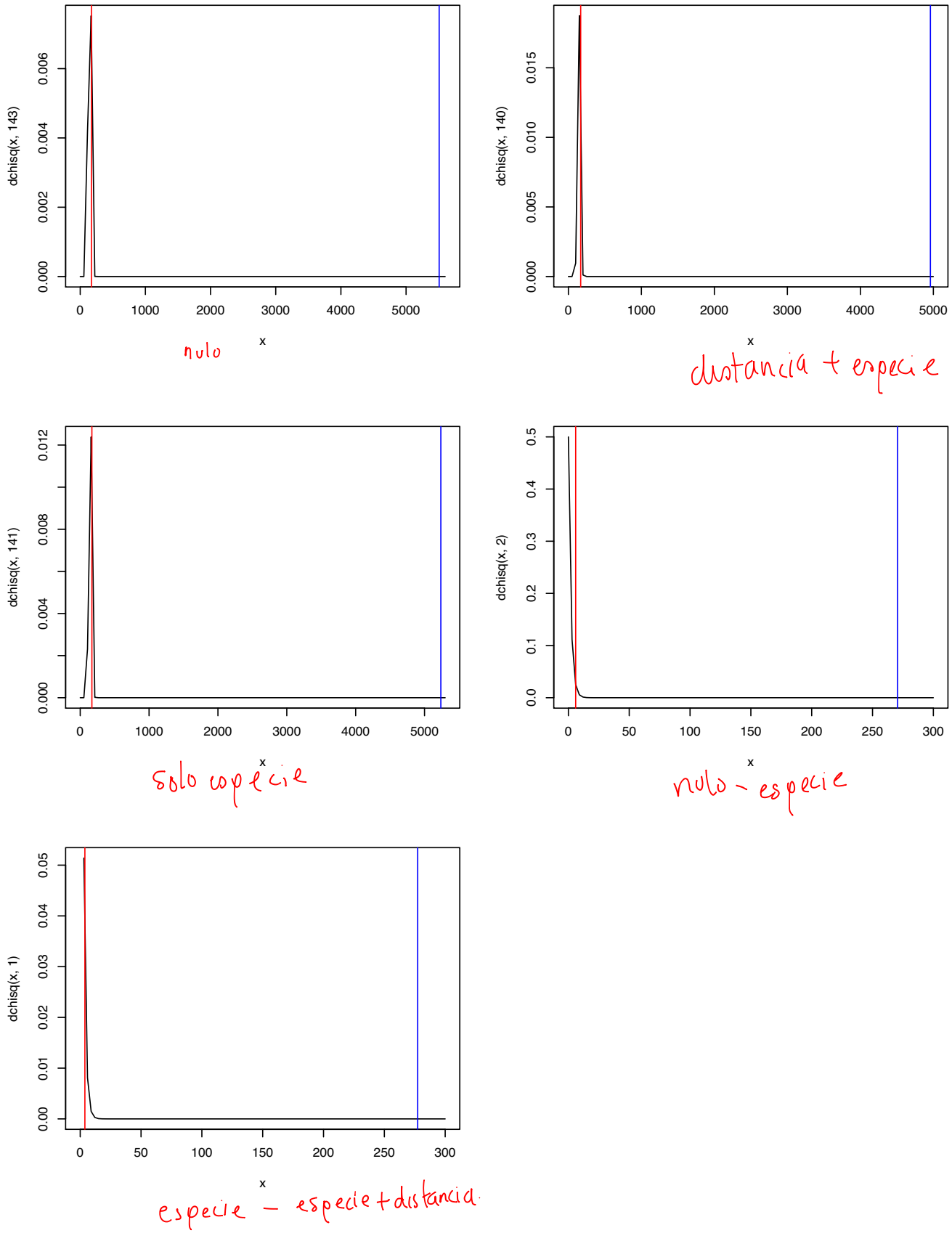

Figura 4: Para cada una de las devianzas (y diferencias de devianzas) calculadas en el modelo 3, se incluye la distribución χ^2 correspondiente (curva negra), el valor crítico para un nivel de significancia del 0.05 (línea roja), y el valor de la devianza (línea azul). De izquierda a derecha y de arriba a abajo: modelo nulo; distancia (m) y especie; solo especie; nulo – especie, y especie – distancia. Se puede observar que los valores observados para todas las devianzas y diferencias de devianzas son muy extremos, lo cual explica también por qué los valores p son tan bajos.

R: Agrupados. Se puede utilizar la devianza de cada variable y compararla con una distribución χ^2 con los grados de libertad correspondientes. Esto ya se ha discutido en el inciso *c)*. Ninguno de los tres modelos presenta un buen ajuste.

g) Interpreta los coeficientes del modelo 3.
R: Para interpretar los coeficientes de forma más natural, se puede aplicar una transformación exponencial

$$logit(\hat{p}) \equiv \log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_1 x_2 + \hat{\beta}_3 x_3 \iff \frac{\hat{p}}{1-\hat{p}} = exp\{\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_1 x_2 + \hat{\beta}_3 x_3\}$$

Así, considerando que en este ejemplo particular x_1 es igual a 1 si la semilla es un frijol y 0 en otro caso, x_2 es igual a 1 si la semilla es de girasol y 0 en otro caso y x_3 es la distancia del camino en la que se dejó la semilla, medida en metros, queda en evidencia que $e^{\hat{\beta}_0} = e^{0.03167} = 1.032177$ es la estimación de la razón de momios de que una semilla de cacahuate dejada a 0 metros del camino sea removida (el modelo estima que casi hay la misma probabilidad de que una semilla con estas características sea removida como de que no). Cuando $x_1 = 1$, $e^{\hat{\beta}_1 x_1} = e^{\hat{\beta}_1}$, y $e^{\hat{\beta}_1} = e^{\hat{\beta}_0 + \hat{\beta}_1 x_1}/e^{\hat{\beta}_0} = e^{\hat{\beta}_1 x_1} e^{-0.77365} = 0.4613262$ es la razón de momios de una semilla de frijol contra una semilla de cacahuate, ambas dejadas a la misma distancia del camino. De esta forma, los momios del frijol son aproximadamente la mitad que los del cacahuate. A través de un razonamiento semejante podemos llegar a la conclusión de que $e^{\hat{\beta}_2} = e^{-0.93717} = 0.3917349$ es la razón de momios de una semilla de girasol contra una de cacahuate, ambas dejadas a la misma distancia del camino. Por tanto, los momios estimados de las semillas de girasol también son aproximadamente un 40 % de los momios del cacahuate. Finalmente, $e^{\hat{\beta}_3} = e^{.0332} = 1.033757$ representa la razón de momios de una semilla del mismo tipo dejada a una distancia x del camino versus una dejada a una distancia $x-1$ (es decir, un metro antes). Dado que la razón de momios es mayor a 1, observamos que los momios de que una semilla sea retirada aumentan (de manera exponencial) conforme se deja a una distancia mayor del camino.

h) Utilizando los coeficientes del modelo 3, ¿Cuál será la probabilidad (estimada) de que una semilla de frijol que se coloca a 10 m del camino sea removida?
R:
$logit(\hat{p}) = 0.03167 - 0.77365 + 0.03320 \times 10 = -0.40998 \iff p = \frac{e^{-0.40998}}{1+e^{-0.40998}} \approx 0.3989169 \approx 0.4$

i) Calcula las probabilidades (estimadas, de ser retiradas) para cada una de las distancias y las semillas evaluadas en este experimento.
R: Siguiendo el procedimiento descrito en el apartado *h* y utilizando los datos del modelo 3 obtenemos la siguiente tabla:

	Distancia(m)		
	0	**15**	**30**
Cacahuate	**0.5079**	**0.6294**	**0.7364**
Frijol	**0.3226**	**0.4393**	**0.5631**
Girasol	**0.2879**	**0.3995**	**0.5223**

j) Realiza un gráfico que muestre las probabilidades calculadas en el punto anterior.
R: El gráfico de probabilidades calculadas en el punto anterior se presenta en la Figura 5.

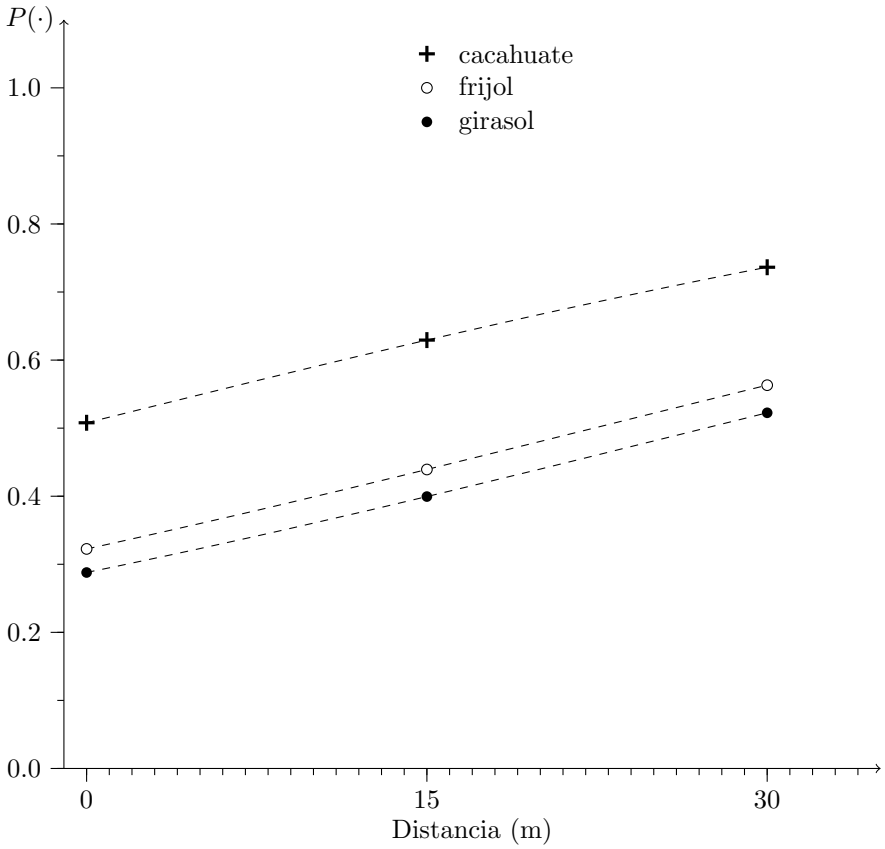

Figura 5: Gráfica de las probabilidades de que una semilla sea removida, según la distancia a la que fue dejada del camino y la especie de la semilla. Solamente se observaron tres distancias, por lo que los íconos representan las condiciones observadas en el experimento y las líneas punteadas son intrapolaciones según el modelo 3.

Referencias

Clopper, C. J., y Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, *26*(4), 404–413.

José, A., y Javier, F. (1993). Razón de posibilidades: una propuesta de traducción de la expresión odds ratio. *Salud Pública de México*, *35*(4), 419–424.

Porta Serra, M. (1990). Traducir or no traducir: ¿es esa la cuestión? *Gaceta Sanitaria*, *4*(16), 38–39.