

Guia 11 Clase

Agustin Muñoz Gonzalez

20/5/2020

Preparamos el entorno

```
rm(list=ls())
```

Las aves parásitas de cría ponen huevos en nidos de otras especies (hospedador), las cuales incuban los huevos y crían al pichón parásito. En un bosque de talas de la provincia de Buenos Aires hay dos especies hospederas que son indistinguibles a simple vista. Una de las principales diferencias entre estas especies radica en el grado de discriminación y remoción de huevos parásitos de sus nidos. Una de las especies es “acceptadora” de huevos parásitos ($Y = 0$), ya que remueve del nido sólo el 30% de los huevos parásitos, mientras que la otra especie es “rechazadora” ($Y = 1$) ya que remueve el 80% de los huevos parásitos presentes en su nido. Además, se sabe que el 90% de los nidos del bosque corresponden a la especie “acceptadora”, mientras que apenas el 10% restante son nidos de la especie “rechazadora”.

Cálculo de probabilidades: Se elige al azar un nido del bosque y se colocan $n = 8$ huevos parásitos. Denotemos con X a la variable aleatoria que indica el número de huevos removidos del nido. Asuma que, en cada nido, la remoción (o no) de los diferentes huevos se realiza de manera independiente.

1. Calcule la probabilidad de que 5 de los huevos parásitos sean removidos del nido si se sabe que el nido parasitado es de la especie aceptadora.

Recopilemos primero la información que tenemos y después respondamos la pregunta. Entendemos por éxito del experimento a que el ave rechaze el huevo parásito. Además tenemos que $Y = 0$ significa que el ave es aceptadora e $Y = 1$ significa rechazadora. Sabemos que

- $P(Y = 0) = 0.9$;
- $P(Y = 1) = 1 - 0.9 = 0.1$;
- $X|Y=0 \sim B(8, 0.3)$;
- $X|Y=1 \sim B(8, 0.8)$.

```
p_A=0.9
```

```
p_R=0.1
```

Ahora sí calculemos lo que nos piden. Queremos saber $P(X = 5|Y = 0)$, es decir queremos la probabilidad de tener 5 éxitos en una variable con distribución binomial. Tenemos que

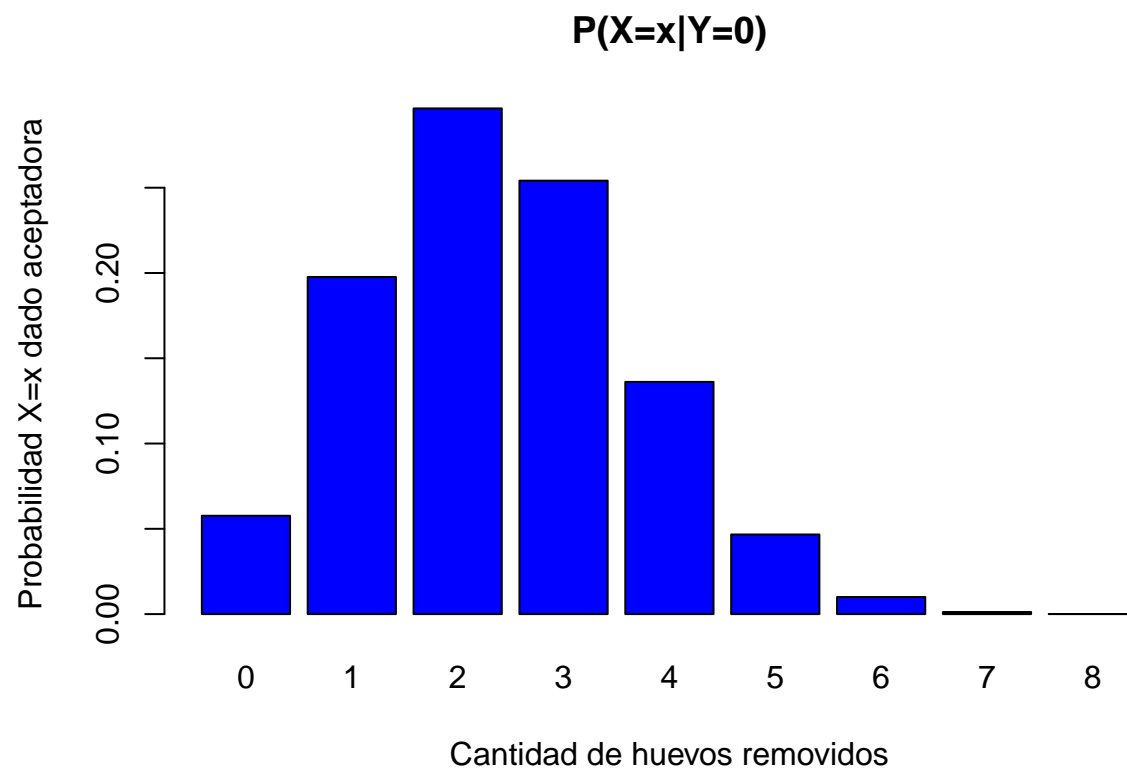
$$P(X = 5|Y = 0) = \binom{8}{5} 0.3^5 0.7^3.$$

```
p_5_dado_A=choose(8,5)*0.3^5*0.7^3
```

2. Calcule la probabilidad de que x de los huevos ajenos sean removidos del nido si se sabe que el nido parasitado es de la especie aceptadora, para $0 \leq x \leq 8$, y represente las probabilidades calculadas en un diagrama de barras. Incluya un título pertinente en el gráfico.

$$P(X = x|Y = 0) = \binom{8}{x} 0.3^x 0.7^{8-x}.$$

```
probas_A=c()
for(i in (0:8)){
  probas_A=c(probas_A,choose(8,i)*0.3^i*0.7^(8-i))
}
barplot(probas_A, names.arg = (0:8),
        xlab='Cantidad de huevos removidos',
        ylab='Probabilidad X=x dado aceptadora',col='blue',
        main='P(X=x|Y=0)')
```



3. Calcule la probabilidad de que 5 de los huevos parásitos sean removidos del nido si se sabe que el nido parasitado es de la especie rechazadora.

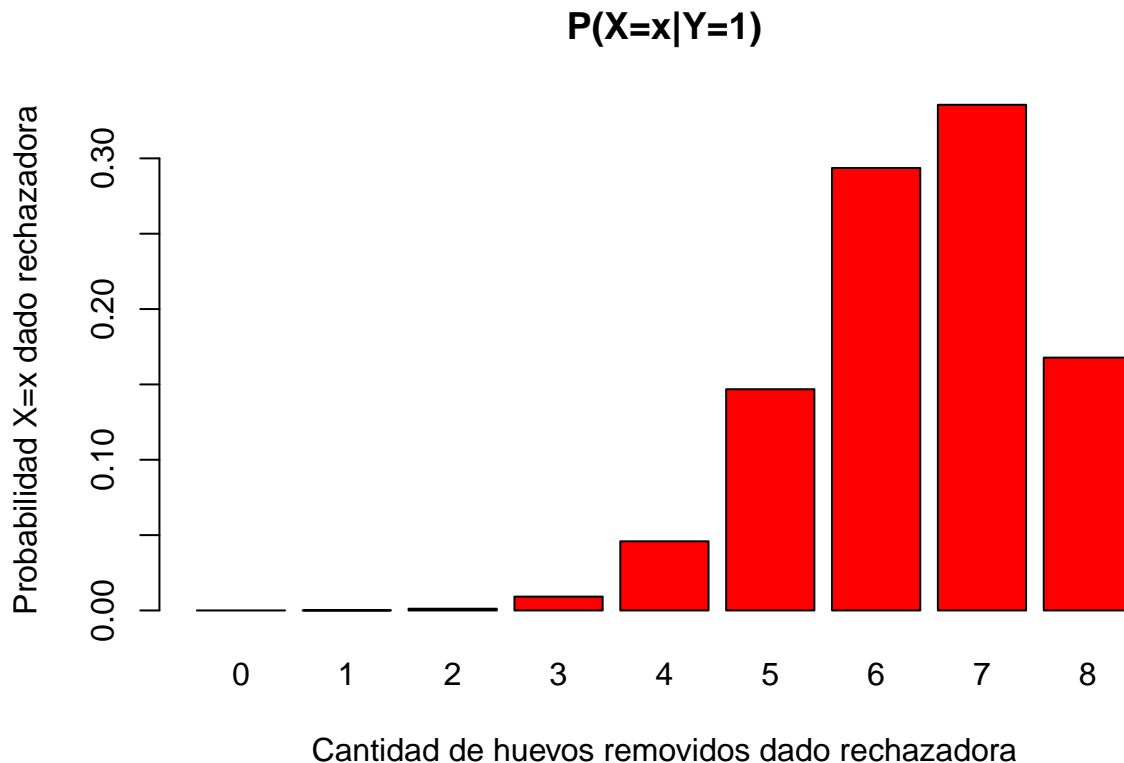
$$P(X = 5|Y = 1) = \binom{8}{5} 0.8^5 0.2^3.$$

```
p_5_dado_R=choose(8,5)*0.8^5*0.2^3
```

4. Calcule la probabilidad de que x de los huevos parásitos sean removidos del nido si se sabe que el nido parasitado es de la especie rechazadora, para $0 \leq x \leq 8$, y representa las probabilidades calculadas en un diagrama de barras. Incluya un título pertinente en el gráfico.

$$P(X = 5|Y = 1) = \binom{8}{5} 0.8^5 0.2^3.$$

```
probas_R=c()
for(i in (0:8)){
  probas_R=c(probas_R,choose(8,i)*0.8^i*0.2^(8-i))
}
barplot(probas_R, names.arg = (0:8),
        xlab='Cantidad de huevos removidos dado rechazadora',
        ylab='Probabilidad X=x dado rechazadora',col='red',
        main='P(X=x|Y=1)')
```



5. Calcule la probabilidad de que 5 de los huevos parásitos sean removidos del nido.

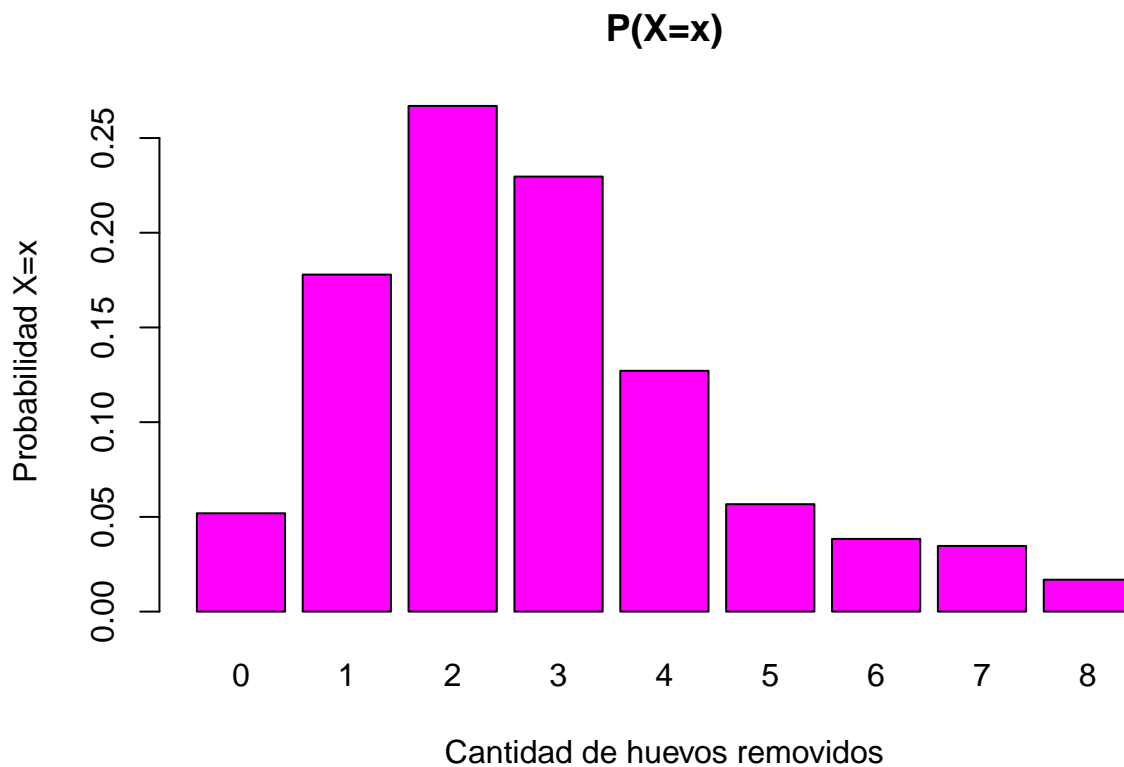
$$P(X = 5) = P(X = 5|Y = 0) * P(Y = 0) + P(X = 5|Y = 1) * P(Y = 1)$$

```
p_5=p_5_dado_A*p_A+p_5_dado_R*p_R
p_5
```

```
## [1] 0.05668796
```

6. Calcule la probabilidad de que x de los huevos parásitos sean removidos del nido, para $0 \leq x \leq 8$, y represente las probabilidades calculadas en un diagrama de barras. Incluya un título pertinente en el gráfico.

```
probas_x=c()
for(i in (0:8)){
  probas_x=c(probas_x,choose(8,i)*(0.3^i*0.7^(8-i)*p_A+0.8^i*0.2^(8-i)*p_R))
}
barplot(probas_x, names.arg = (0:8),
        xlab='Cantidad de huevos removidos',
        ylab='Probabilidad X=x',col='magenta',
        main='P(X=x)')
```



7. Haga la tabla de la función de probabilidad conjunta del vector (X, Y) .

```

probas_conjunta_x_R=probas_conjunta_x_A=c()
for(x in (0:8)){
  proba_x_dado_R=choose(8,x)*0.8^x*0.2^(8-x)
  proba_x_dado_A=choose(8,x)*0.3^x*0.7^(8-x)
  probas_conjunta_x_A=c(probas_conjunta_x_A,proba_x_dado_A*p_A)
  probas_conjunta_x_R=c(probas_conjunta_x_R,proba_x_dado_R*p_R)
}
tabla_conjunta=data.frame(rbind(cbind('Y=0'=probas_conjunta_x_A,'Y=1'=probas_conjunta_x_R,'p_X'=probas_x),
# Esta es una forma de pegarle a una tabla hecha por cols cbind()
# una fila, i.e. rbind(cbind(),c()),
# donde el vector c() tiene que tener la misma longitud que cada fila de cbind().
tabla_conjunta_2=data.frame(cbind(
  rbind('Y=0'=probas_conjunta_x_A,'Y=1'=probas_conjunta_x_R,'p_X'=probas_x),
  c(p_A,p_R,NA)))
names(tabla_conjunta_2)=c((0:8),'p_Y')
#Esta es una forma de pegarle a una tabla hecha de filas rbind(),
# una col, i.e. cbind(rbind(),c()),
#el vector c() tiene que tener la misma longitud que cada col de rbind().

```

OJO LAS PUNTUALES SON $p_{XY}(x, y) = p(X = x, Y = y) = P(X = x|Y = y)P(Y = y)$

Algunos datos: El archivo `depredadosclasificadas` contiene mediciones correspondientes a nidos elegidos al azar en el bosque, registrándose en cada uno de ellos el número de huevos parásitos removidos (primera columna) y la especie del hospedador (segunda columna), determinada mediante un análisis genético. Tenemos entonces observaciones (X, Y) donde X indica la cantidad de huevos removidos del nido, mientras que Y = 0 si el hospedador es aceptador, Y = 1 indica que el hospedador es rechazador.

```
depredados=read.csv('depredadosclasificadas.csv',header = TRUE)
attach(depredados)
```

8. Calcule la proporción de hospedadores aceptadores observados en la muestra. ¿Qué probabilidad está estimando esta proporción?

```
proporcion_Y_A=sum(Y==0)/length(Y)
proporcion_Y_A
```

```
## [1] 0.896
```

Está estimando una probabilidad del 90%.

9. Calcule la proporción de nidos donde se removieron $x = 5$ huevos parásitos entre los hospedadores aceptadores. ¿Qué probabilidad está estimando esta proporción? OJO HAY QUE DIVIDIR NO POR EL TOTAL DE X SINO POR $X|Y=0$ PORQUE EL ESPACIO MUESTRAL CAMBIÓ!

```
proporcion_X_5_MAL=sum(X==5 & Y==0)/length(X)
# Este está mal porque el espacio muestral no es todo S sino S dado Y=0!
proporcion_X_5=sum(X==5 & Y==0)/length(X[Y==0])
# otra forma
mean(X[Y==0]==5)
```

```
## [1] 0.04910714
```

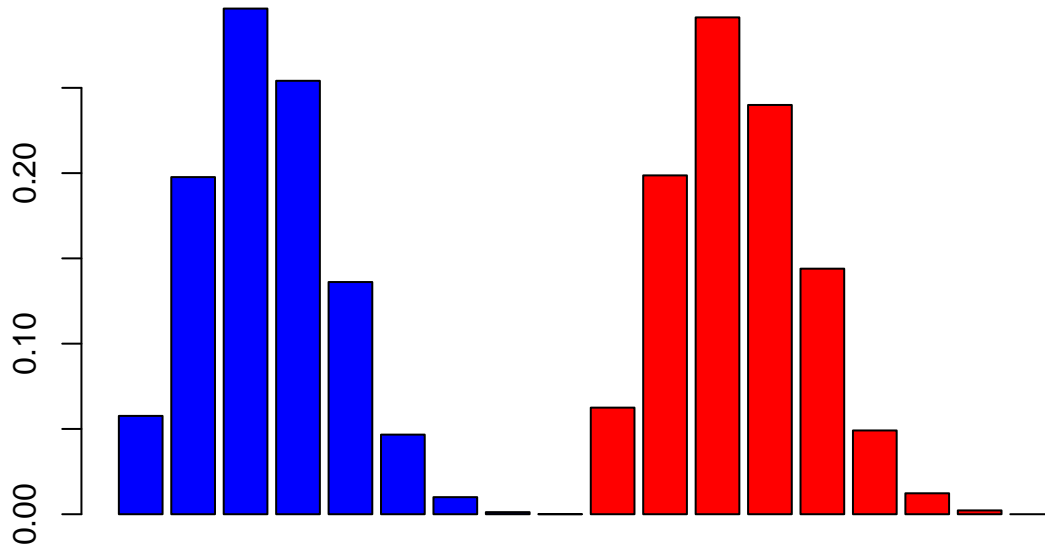
Está estimando una probabilidad del 4.5~5%.

10. Compare la función de probabilidad puntual condicional de X dado $Y = 0$ con la estimada, presentando ambas en un mismo gráfico de barras.

```

probas_A_estimadas=c()
for(i in (0:8)){
  probas_A_estimadas=c(probas_A_estimadas,sum(X==i & Y==0)/length(X[Y==0]))
}
barplot(c(probas_A,probas_A_estimadas),col = rep(c("blue","red"),each=9))

```

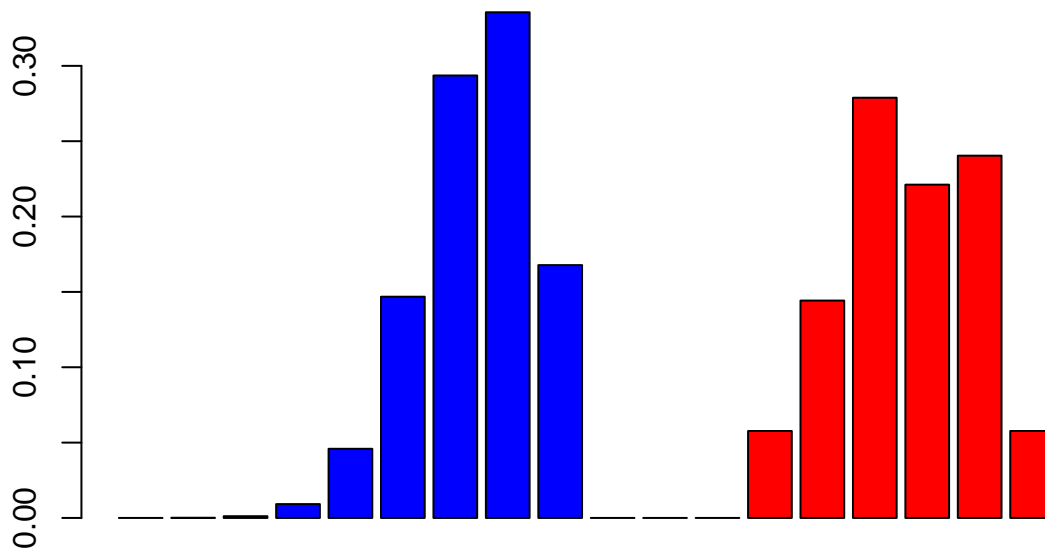


11. Compare la función de probabilidad puntual condicional de X dado $Y = 1$ con la estimada, presentando ambas en un mismo gráfico de barras.

```

probas_R_estimadas=c()
for(i in (0:8)){
  probas_R_estimadas=c(probas_R_estimadas,sum(X==i & Y==1)/length(X[Y==1]))
}
barplot(c(probas_R,probas_R_estimadas),col = rep(c("blue","red"),each=9))

```



12. Hacer la tabla (con dos decimales) de las frecuencias relativas observadas de los pares (x, y) , para $0 \leq x \leq 8$, $0 \leq y \leq 1$. Indique qué valores está estimando con esta tabla.

```
probas_A_estimadas=format(round(probas_A_estimadas, 2), nsmall = 2)
probas_R_estimadas=format(round(probas_R_estimadas, 2), nsmall = 2)
# Nos quedamos con 2 decimales.
datos_frecuencias=matrix(c(cbind(probas_A_estimadas,probas_R_estimadas)),ncol=9,byrow = TRUE)
colnames(datos_frecuencias)=c((0:8))
rownames(datos_frecuencias)=c('y=0','y=1')
tabla_frec_estimadas=data.frame(datos_frecuencias)
tabla_frec_estimadas
```

```
##      X0  X1  X2  X3  X4  X5  X6  X7  X8
## y=0 0.06 0.20 0.29 0.24 0.14 0.05 0.01 0.00 0.00
## y=1 0.00 0.00 0.00 0.06 0.14 0.28 0.22 0.24 0.06
```

Estamos estimando las probabilidades $P(X = x|Y = i)$ para cada x y cada y .