4. Para el siguiente ejercicio es necesario el programa R.

a. Escriba un programa en R que reproduzca las gráficas de las funciones de distribución acumulada y de masa de la distribución uniforme que aparecen en las notas del curso. Las gráficas deben verse similares a las figuras de la Figura 1.

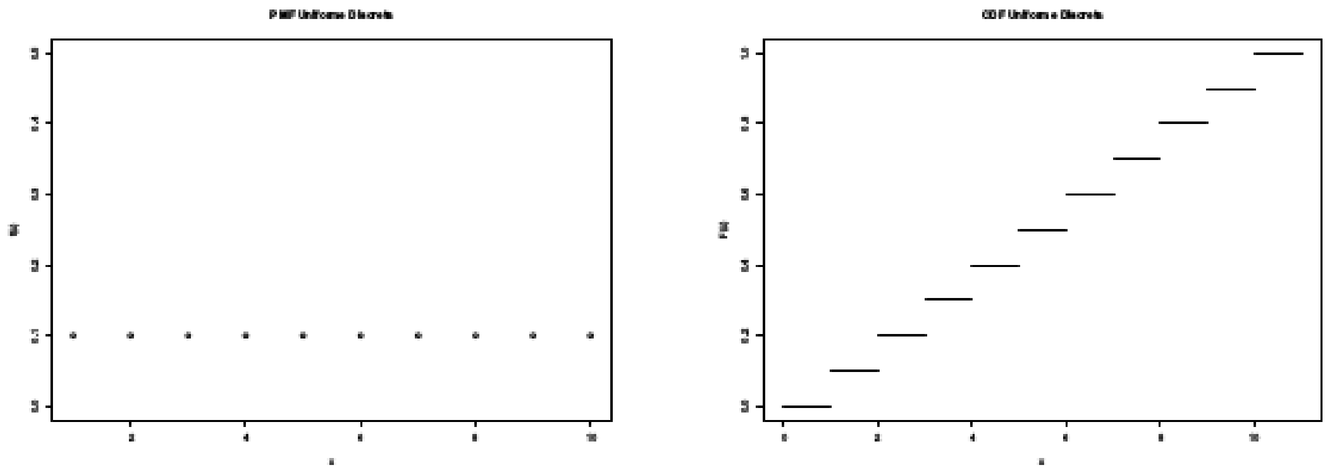


Figura 1: Ejemplo de las funciones de masa y acumulada de la distribución uniforme.

RESPUESTA

Cargamos los paquetes a ocupar:

```
library(tidyverse)
library(gridExtra)
```

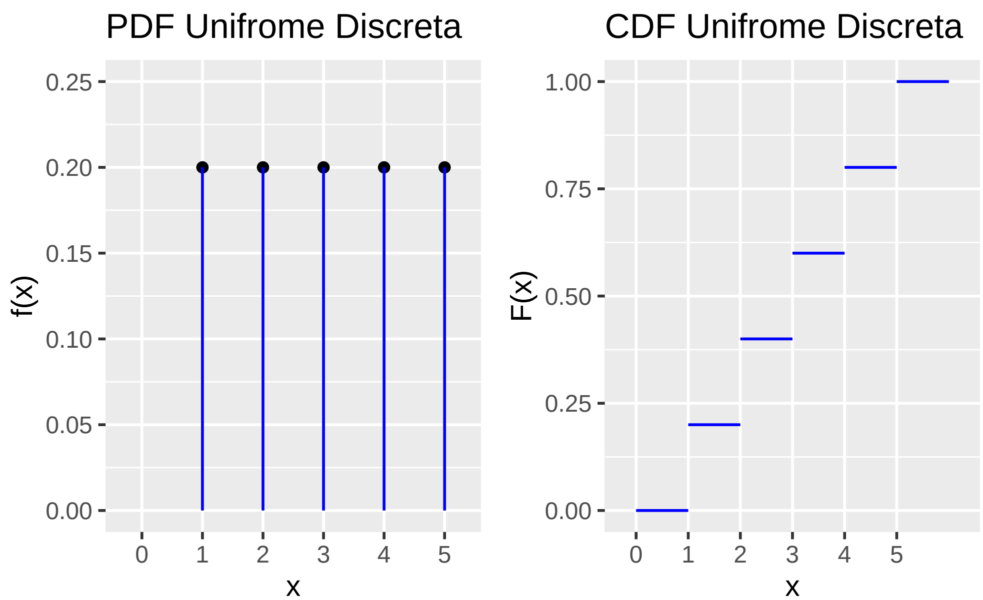
Definimos una función que gráfique la función de probabilidad y la función de probabilidad acumulada para una $U(0, n)$ discreta :

```
grafica_pdf_and_cdf_uniforme <- function(n){ # La función tiene como parametro n.
    data_uniforme <- data_frame(x=seq(1:n)) %>%
      mutate("f_x"=1/n, # Definimos las probabilidades.
             "F_x"= cumsum(f_x)) # Definimos las probabilidades acumaladas.
    data_uniforme[n+1,]<-c(0,NaN,0) # Agregamos el valor inicial

    cdf <- ggplot(data=data_uniforme)+ # CDF
      geom_segment(aes(x=x, xend=x+1, y=F_x, yend=F_x), col="blue")+
      scale_x_discrete(limits=seq(0,n,1))+
      labs(y="F(x)",title="CDF Unifrome Discreta" )+# Graficamos los segementos
      ylim(0,1) # Limites del eje "y".
    pdf <- ggplot(data=data_uniforme)+ # PDF
      geom_point(aes(x,f_x))+
      geom_segment(aes(x=x, xend=x, y=0, yend=f_x), col="blue")+ # Graficamos los segementos.
      scale_x_discrete(limits=seq(0,n,1))+
      ylim(0,1/(n-1))+ # Límites del eje "y".
      labs(x="x", y="f(x)",title = "PDF Unifrome Discreta")
  grid.arrange(pdf, cdf, ncol=2) # Ponemos las gráficas juntas.
  }
```

Mostramos las distribuciones $U(0, 5)$:

```r
grafica_pdf_and_cdf_uniforme(5)
```



b. Lea en la documentación de R, o en cualquier otra fuente de información confiable, la explicación de la función sample(x, size, replace=FALSE, prob=NULL). (No es necesario entregar algo para este ejercicio).

c. Usando la función sample simule una muestra de tamaño 10 000 de la distribución U (1, . . . , 10). Fijando la semilla en 13 (set.seed(13)), muestre los resultados de la simulación en una tabla de frecuencia y calcule la media y la varianza. Sugerencia: Use la función table.

RESPUESTA

```r
set.seed(13) # fijamos la semilla.
muestra <- sample(10, 10000, replace = TRUE) # Muestra aleatoria.
table(muestra) # Tabla de Frecuencias
```

```
## muestra
##    1    2    3    4    5    6    7    8    9   10
## 1015 1001  962 1013 1020  982  991  926 1069 1021
```

```r
mean(muestra) # Media de la muestra simulada.
```

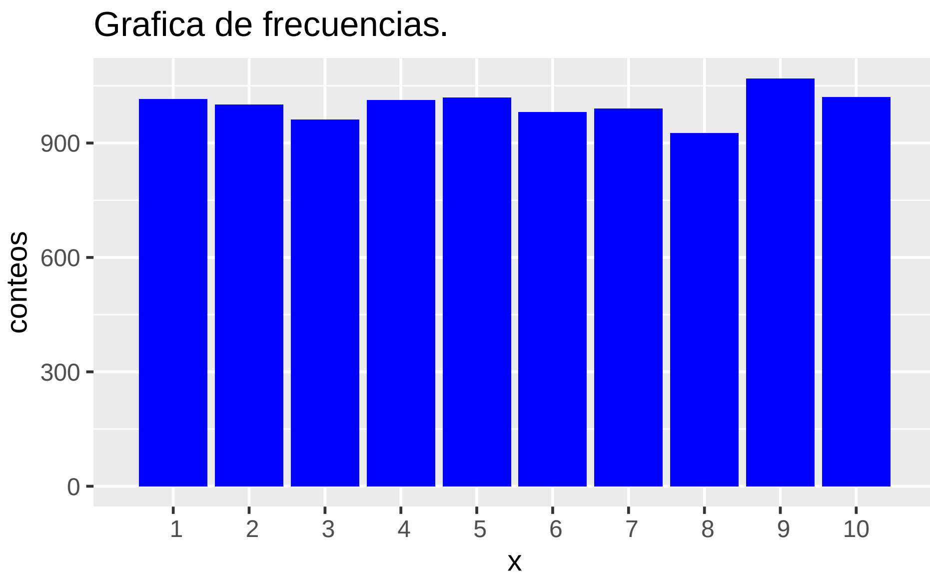
```
## [1] 5.5123
```

```r
var(muestra) # Varianza de la muestra simulada.
```

```
## [1] 8.340283
```

d. Grafique las frecuencias de la simulación anterior.

```r
ggplot(data=data.frame(x=muestra), aes(x))+
  geom_bar(fill="blue")+
  scale_x_discrete(limits =seq(1,10,1))+
  labs(title="Grafica de frecuencias.", y="conteos")
```



5. Para el siguiente ejercicio también necesitamos R.

a. Usando la función sample, simule 10 lanzamientos de una moneda equilibrada y cuente el número de águilas que obtiene. Repita este proceso 10^6 veces y muestre sus primeros 3 resultados. Grafique las frecuencias del número de águilas obtenidas en los 10^6 experimentos. También grafique las proporciones del número de águilas obtenidas.

RESPUESTA

Consideremos, 1: obtener una aguila, 0: obtener sol.

```
resultados_modena <- c(0,1)
sum(sample(resultados_modena, 10, replace = TRUE))
```

```
## [1] 6
```

Ahora repetimos el proceso 10^6 veces.

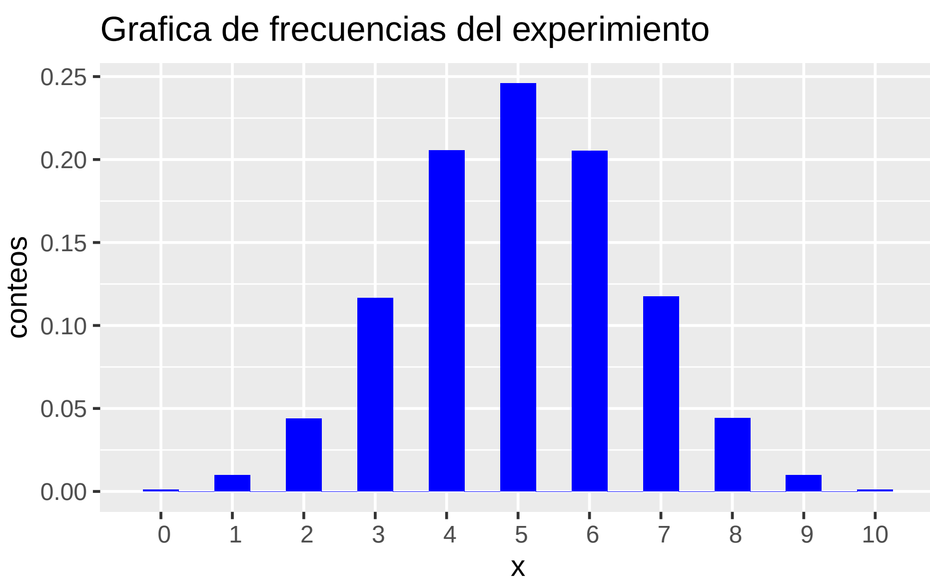
```
simulacion_moneda_equilibrada <- c() # Inicializamos un vector
for(i in 1:10**6) { # Creamos un ciclo para crear las repeticiones.
 simulacion_moneda_equilibrada[i] <-sum(sample(resultados_modena, 10, replace = TRUE))
}
# Mostramos los primeros 3 resultados.
simulacion_moneda_equilibrada[1:3]
```

```
## [1] 3 6 4
```

Graficamos las frecuencias del número de aguilas obtenidas en cada uno de los experimentos.

```
ggplot(data=data.frame(x=simulacion_moneda_equilibrada), aes(x))+
  geom_histogram(bins = 11, binwidth=.5, fill="blue")+
  scale_x_discrete(limits=seq(0,10,1))+
  labs(y="conteos", title="Grafica de frecuencias del experimiento")
```

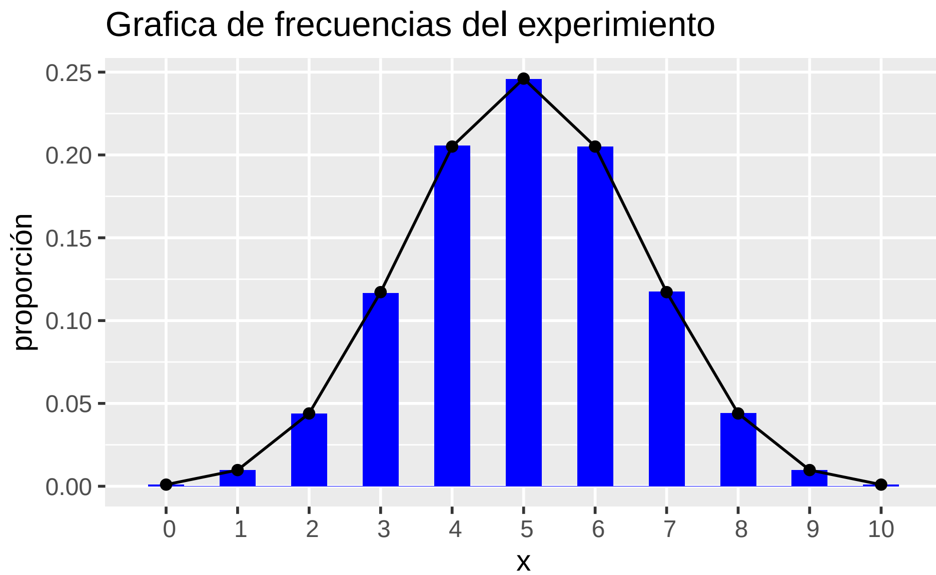

```
ggplot(data=data.frame(x=simulacion_moneda_equilibrada), aes(x))+
  geom_histogram(bins = 11, binwidth=.5, fill="blue", aes(y=stat(count) / sum(count)))+
  scale_x_discrete(limits=seq(0,10,1))+
  labs(y="conteos", title="Grafica de frecuencias del experimiento")
```



b. Usando la función dbinom grafique la función de masa de una distribución B(10, 0.5) sobre la gráfica de las proporciones que hizo en el inciso anterior.

RESPUESTA

```
binomial<-data.frame(x=seq(0,10,1), y=dbinom(x=seq(0,10,1), size = 10,prob = 0.5))
ggplot(data=data.frame(x=simulacion_moneda_equilibrada), aes(x))+
  geom_histogram(bins = 11, binwidth=.5, fill="blue", aes(y=stat(count)/sum(count)))+
  scale_x_discrete(limits=seq(0,10,1))+
  labs(y="proporción", title="Grafica de frecuencias del experimiento")+
  geom_line(data=binomial, aes(x,y))+
  geom_point(data=binomial, aes(x,y))
```



c. Repita los dos incisos anteriores para una moneda desequilibrada que tiene probabilidad $p = 0.3$ de obtener un águila cuando se lanza. ¿Qué observa?

RESPUESTA

Primero consideremos una moneda desequilibrado con probabilidad de 0.3 de obtener águila.

```
lanzamientos_moneda_desequilibrada <- c(rep(0,7),rep(1,3))
```

Simulamos 10 lanzamientos de la6 moneda, y contamos el número de águilas en la simulación:

```
sum(sample(lanzamientos_moneda_desequilibrada, 10, replace = TRUE))
```

```
## [1] 2
```

Ahora realizamos este experimento 10^6 veces.

```
simulacion_monedas_desequilibrada <- c() #inicializamos el vector de resultados.
for(i in 1:10**6) {
simulacion_monedas_desequilibrada[i] <-sum(sample(lanzamientos_moneda_desequilibrada, 10, replace
4}
```

Mostramos los primeros 3 resultados:

```
simulacion_monedas_desequilibrada[1:3]
```

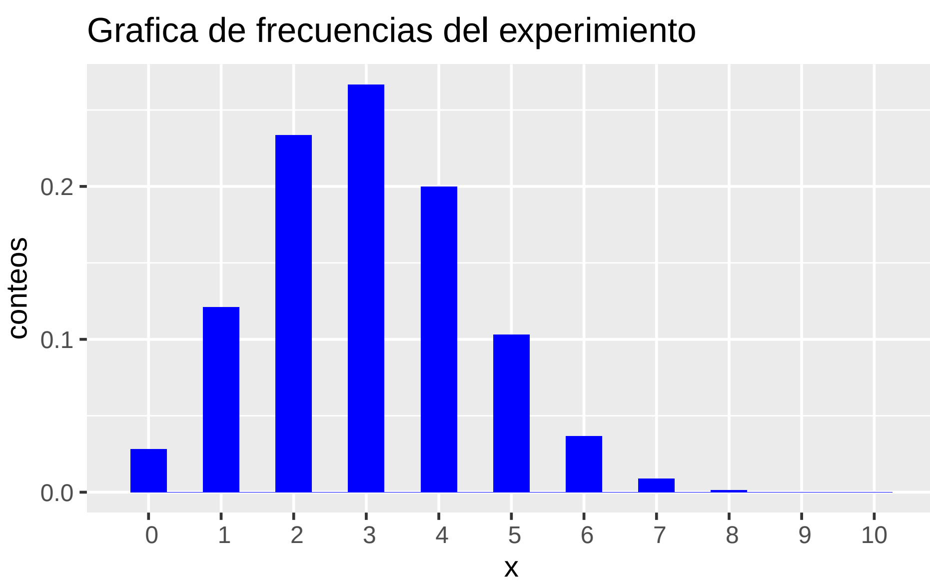
```
## [1] 2 3 4
```

Graficamos las frecuencias del número de aguilas obtenidas en cada uno de los experimentos:

```
ggplot(data=data.frame(x=simulacion_monedas_desequilibrada), aes(x))+
geom_histogram(bins = 11, binwidth=.5, fill="blue")+
scale_x_discrete(limits=seq(0,10,1))+
labs(y="conteos", title="Grafica de frecuencias del experimiento")
```

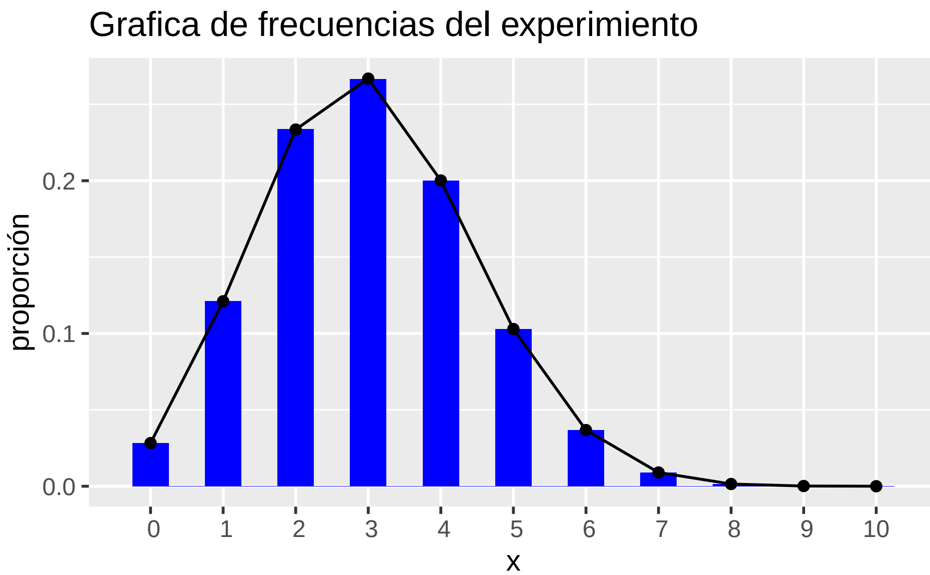
Graficamos las proporciones del número de águilas obtenidas:

```
ggplot(data=data.frame(x=simulacion_monedas_desequilibrada), aes(x))+
geom_histogram(bins = 11, binwidth=.5, fill="blue", aes(y=stat(count)/sum(count)))+
scale_x_discrete(limits=seq(0,10,1))+
labs(y="conteos", title="Grafica de frecuencias del experimiento")
```



Usando la función dbinom graficamos la función de masa de una distribución $B(10, 0.5)$ sobre la gráfica anterior:

```
binomial<-data.frame(x=seq(0,10,1), y=dbinom(x=seq(0,10,1), size = 10,prob = 0.3))
ggplot(data=data.frame(x=simulacion_monedas_desequilibrada), aes(x))+
geom_histogram(bins = 11, binwidth=.5, fill="blue", aes(y=stat(count)/sum(count)))+
scale_x_discrete(limits=seq(0,10,1))+
labs(y="proporción", title="Grafica de frecuencias del experimiento")+
geom_line(data=binomial, aes(x,y))+
geom_point(data=binomial, aes(x,y))
```



6. Una urna contiene 46 bolas grises y 49 bolas blancas. Usando la función sample en R, simule la extracción sin reemplazamiento de 20 de estas bolas y cuente el número de bolas grises que obtuvo. Repita este proceso 10^6 veces y grafique las frecuencias de bolas grises obtenidas en cada experimento. ¿Cuál es la probabilidad de que al extraer 20 bolas de la urna 5 de ellas sean grises? También grafique la proporción de bolas grises obtenidas en los experiementos anteriores y sobre esta figura añada la correspondiente función de masa de la distristibución Hipergeometrica asociada al experimento total.

RESPUESTA

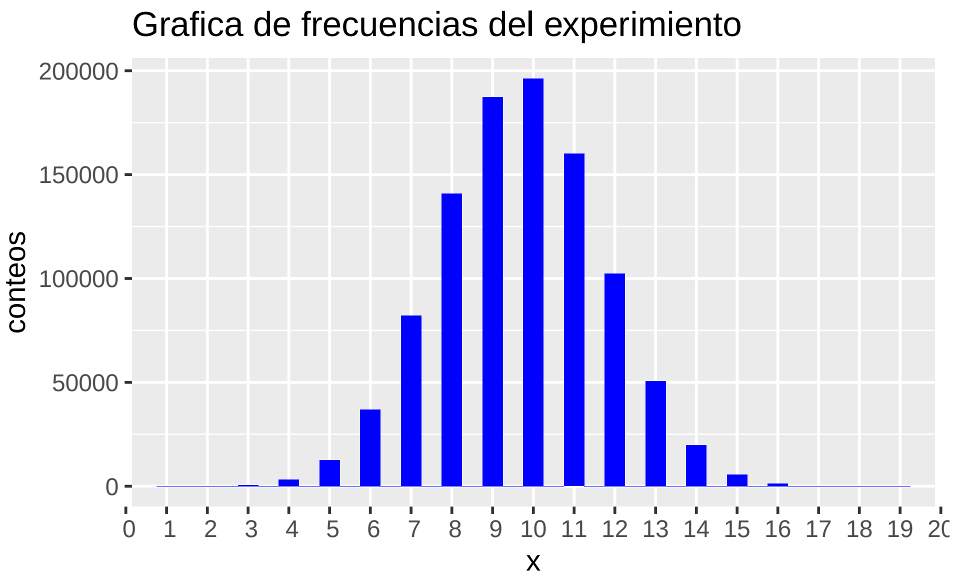
Sea 1: una bola gris, y 0: una bola blanca.

```
urna <- c(rep(1,46), rep(0,49)) # Creamos la urna con las bolas.
sum(sample(urna, 20, replace=FALSE)) # Contamos cuantas bolas grises hay en la muestra.
```

```
## [1] 9
```

```
simulacion_bolas <- c() #Inicializamos
for(i in 1:10**6) { # Repetimos el experimento de extraer una muestra.
 simulacion_bolas[i] <-sum(sample(urna, 20, replace=FALSE))
}
```

```
ggplot(data=data.frame(x=simulacion_bolas), aes(x))+
  geom_histogram(bins = 19, binwidth=.5, fill="blue")+
  scale_x_discrete(limits=seq(0,20,1))+
  labs(y="conteos", title="Grafica de frecuencias del experimiento")
```



Si X es el número es de bolas grises que se obtienen en la extracción sin reemplazamiento de 20 bolas de la urna definida en el problema, entonces podemos decir que $X \sim Hyper(n = 20, M = 46, N = 95)$. Por lo tanto, la porbabilidad de que se extraigan 5 bolas grises es:

$$f(x = 5) = \frac{\binom{46}{5}\binom{49}{15}}{\binom{95}{20}} = 0.01261935.$$

Graficamos la proporción de bolas grises:

```
hyper<-data.frame(x=seq(0,20,1), y=dhyper(x=seq(0,20,1), m = 46, n=49, k=20))
ggplot(data=data.frame(x=simulacion_bolas), aes(x))+
  geom_histogram(bins = 21, binwidth=.5, fill="blue", aes(y=stat(count)/sum(count)))+
  scale_x_discrete(limits=seq(0,20,1))+
  labs(y="proporción", title="Grafica de frecuencias del experimiento")+
  geom_line(data=hyper, aes(x,y))+
  geom_point(data=hyper, aes(x,y))
```