Distribución geométrica (Variables aleatorias discretas)

Probabilidad I

2021-1

César Augusto Pérez Rosas

Facultad de Ciencias

1 de diciembre de 2020

Variable aleatoria Geométrica

- Suponga que se realizan una infinidad de ensayos independientes del tipo Bernoulli con propbabilidad de éxito $= p$

Variable aleatoria Geométrica

- Suponga que se realizan una infinidad de ensayos independientes del tipo Bernoulli con propbabilidad de éxito $= p$
- Para cada realización del experimento se define la v.a. X como el número de intentos antes del primer éxito

Variable aleatoria Geométrica

- Suponga que se realizan una infinidad de ensayos independientes del tipo Bernoulli con propbabilidad de éxito $= p$
- Para cada realización del experimento se define la v.a. X como el número de intentos antes del primer éxito
- Los posibles valores que puede tomar esta variable aleatoria son:

Variable aleatoria Geométrica

- Suponga que se realizan una infinidad de ensayos independientes del tipo Bernoulli con propbabilidad de éxito $= p$
- Para cada realización del experimento se define la v.a. X como el número de intentos antes del primer éxito
- Los posibles valores que puede tomar esta variable aleatoria son:
- Por ejemplo:

Definición(Variable aleatoria Geométrica)

Una v.a. X se dice que tiene distribución Geométrica con parámetro p, $0 < p < 1$, denotada como $X \sim Geo(p)$, si su función de masa de probabilidad está dada por la siguiente expresión:

$$f(x) = \begin{cases} (1-p)^x p & \text{para } x = 0, 1, 2, \dots \\ 0 & \text{en otro caso} \end{cases}$$

Definición(Variable aleatoria Geométrica)

Una v.a. X se dice que tiene distribución Geométrica con parámetro p, $0 < p < 1$, denotada como $X \sim Geo(p)$, si su función de masa de probabilidad está dada por la siguiente expresión:

$$f(x) = \begin{cases} (1-p)^x p & \text{para } x = 0, 1, 2, \ldots \\ 0 & \text{en otro caso} \end{cases}$$

Series geométricas

- $\sum_{k=0}^{n} ar^k = a\frac{1-r^{n+1}}{1-r}$ para $r \neq 1$
- $\sum_{n=0}^{\infty} ar^n = \frac{1}{1-r}$ si $|r| < 1$

Momentos

Para una variable aleatoria X tal que $X \sim Geo(p)$ se tiene que:

$$\mathbb{E}(X) = \frac{1-p}{p}$$

$$\mathsf{Var}(X) = \frac{1-p}{p}$$

$$m_X(t) = \frac{p}{1-(1-p)e^t}$$

Función de masa de probabilidad en R

Se puede encontrar la $\mathbb{P}(X = x)$ para una v.a. con distribución geométrica usando la función `dgeom()`, por ejemplo:

$\mathbb{P}(X = x)$ para $0 \leq x \leq 15$, $x \in \mathbb{N}$, donde $X \sim Geo(0.2)$

```
dgeom(0:15,.2)

##  [1] 0.200000000 0.160000000 0.128000000 0.102400000 0.081920000 0.065536000
##  [7] 0.052428800 0.041943040 0.033554432 0.026843546 0.021474836 0.017179869
## [13] 0.013743895 0.010995116 0.008796093 0.007036874
```

$\mathbb{P}(X = 6)$ donde $X \sim Geo(0.2)$

```
dgeom(6,p=.2)

## [1] 0.0524288
```

Teorema(Propiedad de perdida de memoria)

Sea X una variable aleatoria tal que $X \sim Geo(p)$, y sea $k \geq 0$. Entonces para cada entero $t \geq 0$,

$$\mathbb{P}(X = k + t | X \geq k) = \mathbb{P}(X = t)$$

Función de distribución acumulada en R

Se puede encontrar la probabilidad acumulada para una v.a. con distribución Geométrica usando la función `pgeom()`, por ejemplo:
$\mathbb{P}(X \leq x)$ para $0 \leq x \leq 15$ donde $X \sim Geo(0.2)$

```
pgeom(0:15,p=.2)

##  [1] 0.2000000 0.3600000 0.4880000 0.5904000 0.6723200 0.7378560 0.7902848
##  [8] 0.8322278 0.8657823 0.8926258 0.9141007 0.9312805 0.9450244 0.9560195
## [15] 0.9648156 0.9718525
```

$\mathbb{P}(X \leq 10)$ donde $X \sim Geo(0.2)$

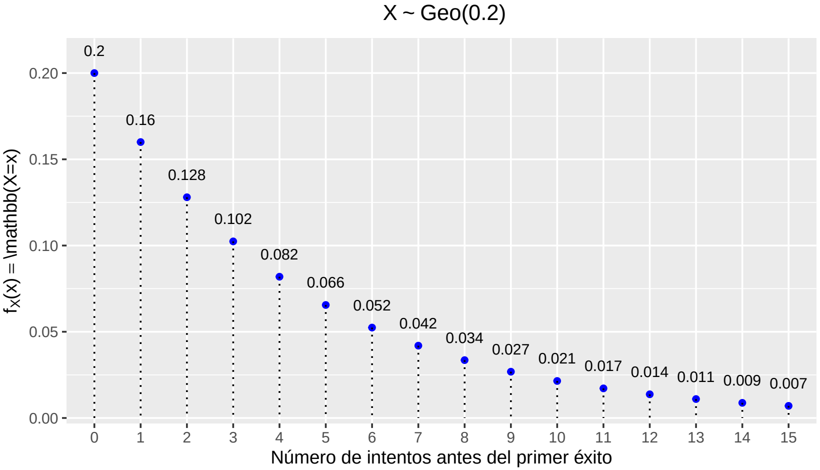
```
pgeom(10,p=.2)

## [1] 0.9141007
```

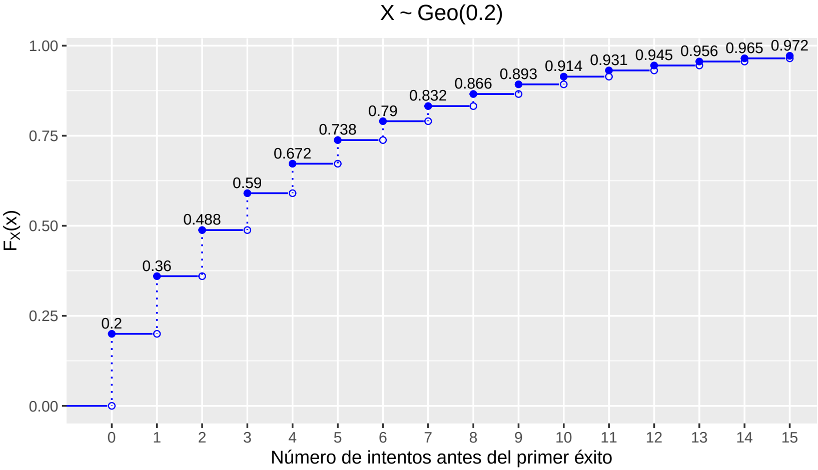
Ejemplo

Una urna contiene N bolas blancas y M bolas negras. Se seleccionan bolas aleatoriamente con reemplazo hasta que se obtiene la primer bola negra. ¿Cuál es la probabilidad de sacar excatamente n bolas?

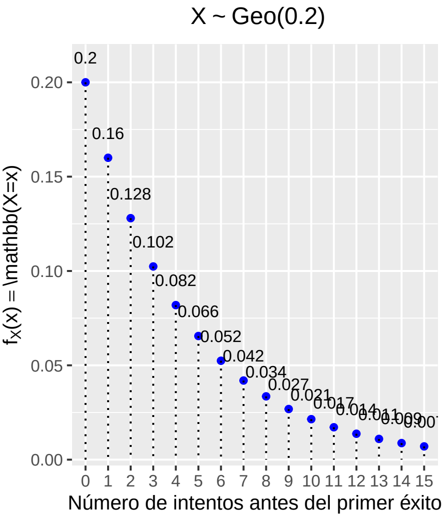
$f(x) = \mathbb{P}(X = x)$

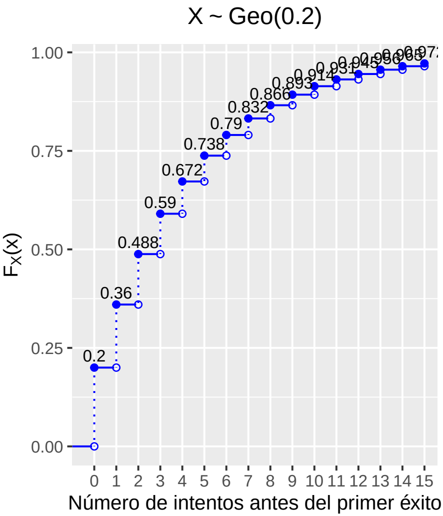


$F(x) = \mathbb{P}(X \leq x)$



Gráficas $F_X(x)$ y $f_X(x)$





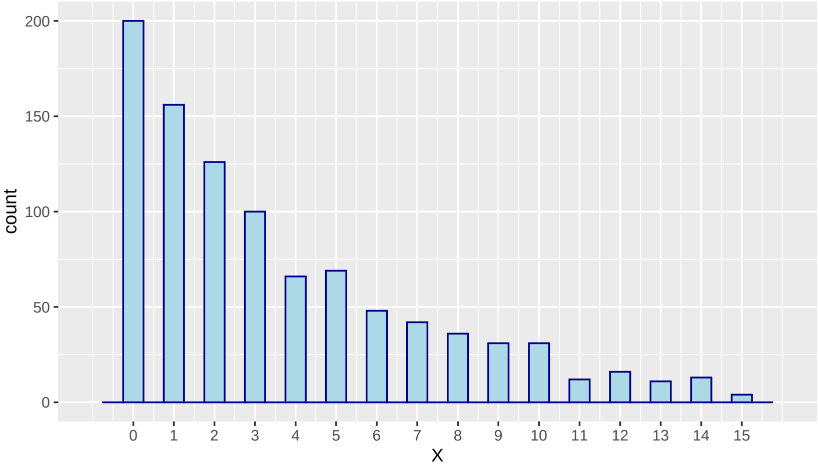
Simulación de valores de la distribución Geométrica en R

Para simular valores de la distribución geométrica en R se puede hacer uso de la función `rgeom(valores a simular,`λ `)`, por ejemplo:

```
#Simulación de 150 valores de la distribución Geométrica(p=.2)
rgeom(150,p=0.2)

##   [1] 11  0  1  5  5  6  3  7  0  0  1  5  1  0  8  1  7  5  2  9  9  2  3  2  8
##  [26]  0  6  2 12  7  4  4  3  1  0  1  1 11  2 11  1  3  0  1  2  8  5 11  0  0
##  [51]  3  3 10  5  0  0  7 10  0 10  4  2 15  7  1 12 19  1  1  4  4 20  4  0  2
##  [76]  0  3  9  5  3  1  1  4  4  6  4  1  6  4  3 26 11  0  2  7  3  2 14  4  1
## [101]  3  3  3  1  3  8  0  2  1  3  0 15  5  1  1 10 13  1  3  0  6  3  4  0  0
## [126]  3  1  2  5 14  0  4  6  3  2  0  0  0  9  5  7 11  5  0  9  7  2  4  3  0
```

```r
simula_geom<-rgeom(1000,p=0.2)
ggplot(mapping = aes(simula_geom))+
  geom_histogram(binwidth = 0.5,color="darkblue",fill="lightblue")+
  scale_x_continuous(name = "X",breaks = seq(0, 15, 1),
                     limits=c(-1,16))
```



```r
table(simula_geom)

## simula_geom
##   0   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15
## 200 156 126 100  66  69  48  42  36  31  31  12  16  11  13   4
##  20  21  22  23  25  26  29
##   2   5   1   2   1   1   1

cumsum(simula_geom)[1000]/length(simula_geom)

## [1] 4.226
```

```r
library(kableExtra)
library(latex2exp)
kable(data.frame("x"=c(0:10),"y"=dgeom(0:10,0.2),"z"=pgeom(0:10,0.2)),
col.names = c("$x$","$\\mathbb{P}(X)=x$","$\\mathbb{P}(X)\\leq x$"),
escape = F,booktabs = T)%>%
  kable_styling(latex_options = c("striped"),font_size = 8)
```

x	$\mathbb{P}(X) = x$	$\mathbb{P}(X) \leq x$
0	0.2000000	0.2000000
1	0.1600000	0.3600000
2	0.1280000	0.4880000
3	0.1024000	0.5904000
4	0.0819200	0.6723200
5	0.0655360	0.7378560
6	0.0524288	0.7902848
7	0.0419430	0.8322278
8	0.0335544	0.8657823
9	0.0268435	0.8926258
10	0.0214748	0.9141007

Histograma para distintos valores de p(5000 simulaciones)