

Distribución Binomial

Maldonado Silva Nancy Magali

Introducción

La distribución probabilística binomial es una distribución de probabilidades discreta. Una característica de esta distribución binomial (o binómica) es que se ocupa de experimentos donde cada resultado puede tomar sólo una de dos formas. Por ejemplo, la respuesta a una pregunta del tipo verdadero o falso es precisamente “verdadero” o “falso”. Cada resultado es mutuamente excluyente, lo que significa en este caso, que la respuesta a una pregunta del tipo verdadero no puede ser correcta y estar equivocada al mismo tiempo.

Una forma común de denotar los dos resultados es como “éxito” y “fracaso”. Por ejemplo, si adivinara la respuesta correcta a una pregunta de verdadero o falso el resultado se clasificará como éxito. En caso contrario, es un fracaso.

Una segunda característica de una distribución binomial es que los datos recopilados son resultados de conteos. Esta es una razón por la que la distribución binomial se clasifica como distribución discreta.

Una tercera peculiaridad de esta distribución es que la probabilidad de un éxito permanece igual de un ensayo a otro.

Una cuarta característica de una distribución binomial es que un ensayo es independiente de cualquier otro. En realidad, esto es lo mismo que decir que no existe un patrón rítmico con respecto a los resultados (Georgecanavos 1988).

La suma de k variables independientes e idénticamente distribuidas $(0, 1)$ tiene una distribución binomial (Altham 2018).

El presente manual expone algunos ejemplos de aplicación mediante un paquete de R (R Core Team 2022) construido que se le ha denominado `DistribucionBinomial`. El código de fuente del paquete está alojado en la siguiente dirección de GitHub <https://github.com/mNancyMa/DistribucionBinomial.git>.

La instalación del paquete en cualquier computadora se realiza ejecutando en la consola de R lo siguiente:

```
devtools::install_github("mNancyMa/DistribucionBinomial")
```

Objetivos

1. Predicción de eventos: la distribución binomial permite predecir la probabilidad de que ocurra un evento determinado en un grupo de animales. Por ejemplo, se puede usar para estimar la probabilidad de que un cierto número de animales de un rebaño sean de un determinado sexo.
2. Evaluación de frecuencia: Permite determinar si los resultados observados son consistentes con lo que se esperaría al azar, o si hay alguna influencia adicional que pueda surgir de los resultados.
3. Estimación de parámetros: se utilizará para estimar parámetros importantes en la zootecnia, como la tasa de éxito (probabilidad de ocurrencias de un evento) o la proporción de animales con una característica determinada.

Modelo estadístico

La distribución binomial tiene como modelo la siguiente ecuación:

$$P = (n!x)p^x(1 - p)^{n-x} \quad (1)$$

$$(n!x) = n!/(x!(n - x)!) \quad (2)$$

Donde

- n Representa el número de experimentos
- x Indica el número de éxitos
- p Probabilidad de éxito
- $1 - p$ Probabilidad de fracaso
- $(n!x)$ Combinatoria de n elementos tomados de x en x

Ejemplo de aplicación del paquete

La aplicación de la distribución binomial es de gran relevancia en la zootecnia ya que podemos modelar y analizar eventos discretos de éxito o fracaso en una serie de ensayos independientes. Se puede aplicar en los siguientes casos:

- **Análisis de datos de experimentos:** Experimentos para evaluar diferentes variables y medidas en animales. Por ejemplo será útil para evaluar la eficacia de ciertos tratamientos o intervenciones en los animales.
- **Estimación de parámetros genéticos:** En la cría de animales, es fundamental comprender y estimar los parámetros genéticos de diferentes características. Se puede usar para estimar la heredabilidad, que es la proporción de la variación fenotípica debida a la variación genética. Esto ayuda a los zootecnistas a seleccionar y mejorar características específicas en la crianza de animales.
- **Modelado de tasas de mortalidad y supervivencia:** Será esencial para comprender y predecir en diferentes poblaciones de animales. Se puede modelar y analizar estas tasas, lo que proporciona información valiosa para la gestión y el diseño de programas de crianza y conservación.
- **Evaluación de la probabilidad de eventos sanitarios:** En la salud animal, es importante evaluar y predecir la probabilidad de eventos sanitarios, como brotes de enfermedades. Esto permitirá estimar la probabilidad de que un cierto número de animales se vean afectados por una enfermedad en una población determinada, lo que es fundamental para la implementación de medidas de control y prevención.

Ejemplo 1

En una granja lechera, se realizó un muestro de prueba de mastitis, para determinar la cantidad de vacas que tenían este problema. Se examinaron un total de 15 vacas. A las vacas que no resultaran afectadas por la mastitis se les consideró fracaso, que se denotaba con un 0, caso contrario con las vacas que padecían de mastitis, estas fueron contadas como éxitos durante el experimento, denotándolos con el número 1. Terminado el muestreo (experimento), se tuvo un total de 8 éxitos, es decir, 8 de las 15 vacas resultaron positivas a la prueba de mastitis. Para el experimento, se ocupó una probabilidad del 0.5

```
directorio <- "C:\\Users\\CICEM\\OneDrive\\Escritorio\\ProbabilidadBinomial\\Prueba\\Datos
datos <- read.csv(directorio)
y <- datos[, 2]
y
```

```
[1] 0 0 1 1 0 1 0 1 0 1 1 1 0 1 0
```

```
knitr::kable(datos)
```

numvacas	y
1	0
2	0
3	1
4	1
5	0
6	1
7	0
8	1
9	0
10	1
11	1
12	1
13	0
14	1
15	0

Para calcular la distribución binomial primero cargamos el paquete **ProbBin** mediante la siguiente instrucción:

```
library(ProbBinom)
```

Posteriormente, se obtiene el cálculo de la probabilidad binomial, junto con la acumulativa llamando la función **Funcbin**, el primer argumento que recibe la función es *y*, que en este caso es la cantidad de experimentos y como siguiente argumento de la función es la probabilidad, que en este caso es de 0.5.

```
resultado <- FuncBin(y,p=0.5)
```

Número de éxitos y fracasos en Y

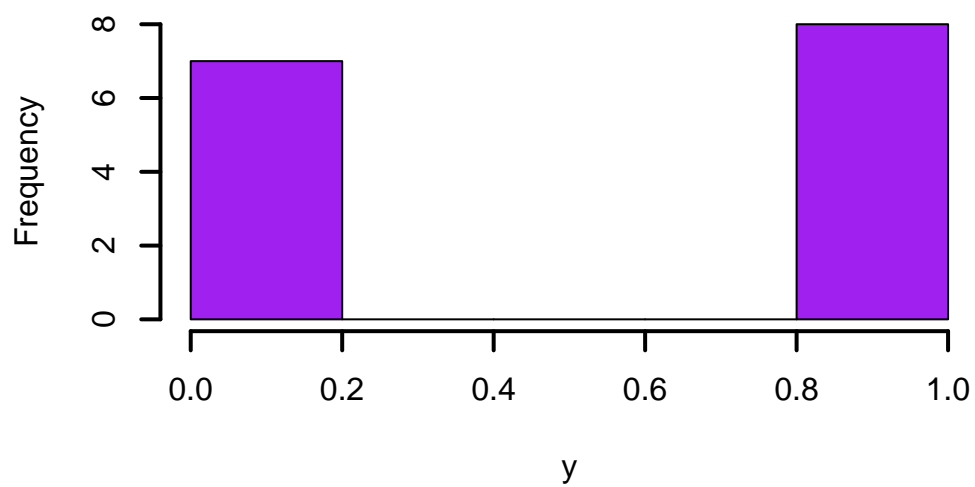
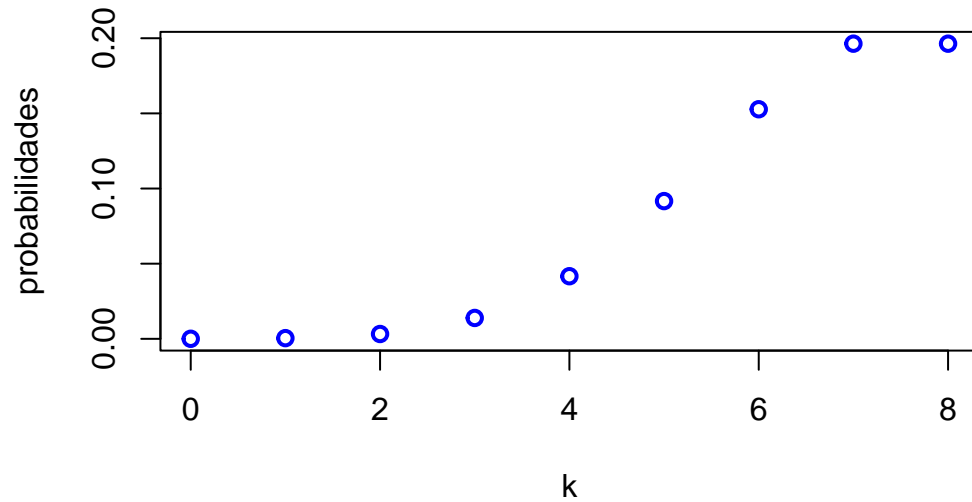
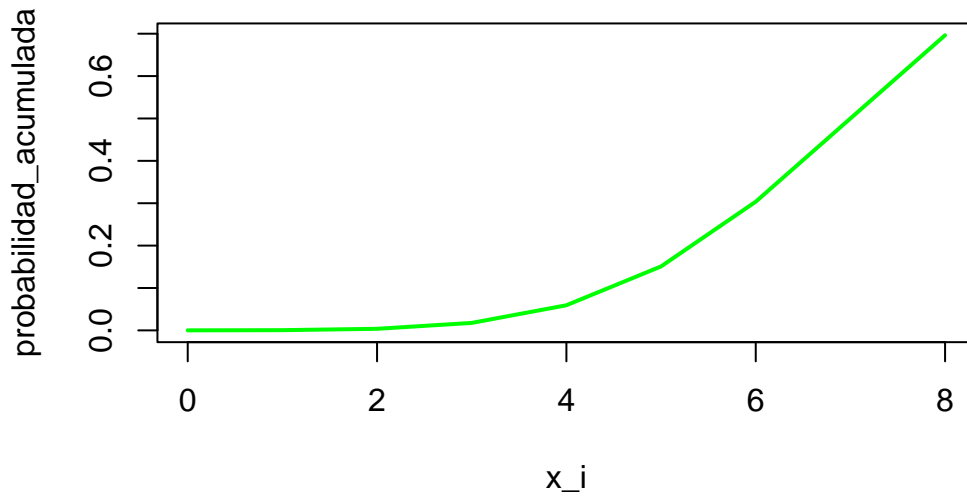


Gráfico De Probabilidades



Probabilidad Acumulada



resultado

[[1]]

x	Coefficiente	Probabilidad	x_i	coeficiente_i	Probabilidades	Acumulada
1	8	6435	0.1963806	0	1	3.051758e-05
2	8	6435	0.1963806	1	15	4.577637e-04
3	8	6435	0.1963806	2	105	3.204346e-03
4	8	6435	0.1963806	3	455	1.388550e-02
5	8	6435	0.1963806	4	1365	4.165649e-02
6	8	6435	0.1963806	5	3003	9.164429e-02
7	8	6435	0.1963806	6	5005	1.527405e-01
8	8	6435	0.1963806	7	6435	1.963806e-01
9	8	6435	0.1963806	8	6435	1.963806e-01

[[2]]

\$breaks

[1] 0.0 0.2 0.4 0.6 0.8 1.0

\$counts

[1] 7 0 0 0 8

\$density

```
[1] 2.333333 0.000000 0.000000 0.000000 2.666667
```

```
$mids
```

```
[1] 0.1 0.3 0.5 0.7 0.9
```

```
$xname
```

```
[1] "y"
```

```
$equidist
```

```
[1] TRUE
```

```
attr(,"class")
```

```
[1] "histogram"
```

```
[[3]]
```

```
NULL
```

Ejemplo 2

Se realiza un experimento utilizando la distribución binomial para saber la cantidad de ovejas que tienen garrapatas. Para este experimento, hallar garrapatas sería el éxito que se denota con el número “1”, y para las ovejas sin garrapatas sería fracaso o “0”. Para el experimento se contó con 25 ovejas, de las cuales 12 resultaron exitosos. Para este experimento se utilizó una probabilidad de éxito del 0.45.

En este caso, los datos no se introducen manualmente en R, mediante un `data.frame`.

```
df <- data.frame(  
  núm_ovejas=c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25),  
  y=c(0,0,1,1,1,0,0,0,1,1,0,0,1,0,1,1,1,0,1,0,1,0,0,0,1)  
)  
y <- df[,2]  
y
```

```
[1] 0 0 1 1 1 0 0 0 1 1 0 0 1 0 1 1 1 0 1 0 1 0 0 0 1
```

```
knitr::kable(df)
```

núm_ovejas	y
1	0

núm_ovejas	y
2	0
3	1
4	1
5	1
6	0
7	0
8	0
9	1
10	1
11	0
12	0
13	1
14	0
15	1
16	1
17	1
18	0
19	1
20	0
21	1
22	0
23	0
24	0
25	1

Después se procede a realizar los cálculos de la distribución binomial, cargando el paquete ProbBinom

```
library(ProbBinom)
```

Se carga la función para obtener los resultados

```
garrapatas <- FuncBin(y, 0.45)
```


Número de éxitos y fracasos en Y

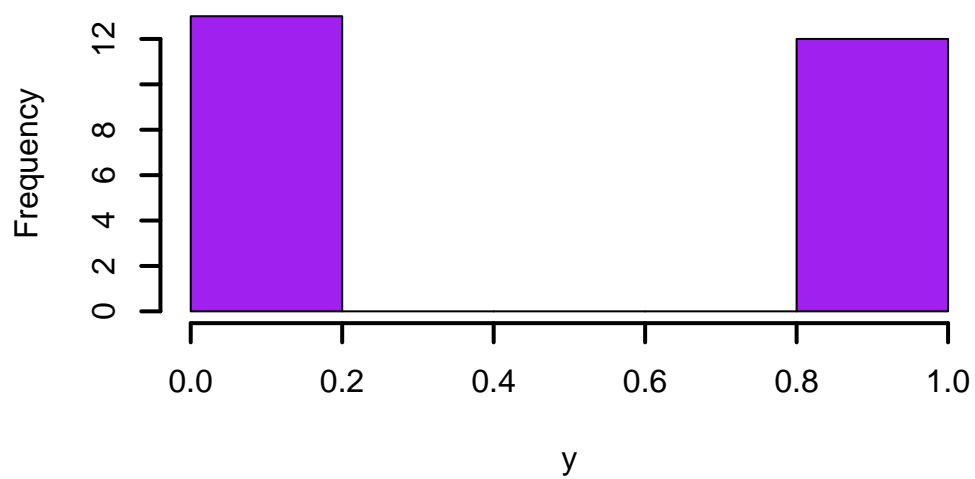
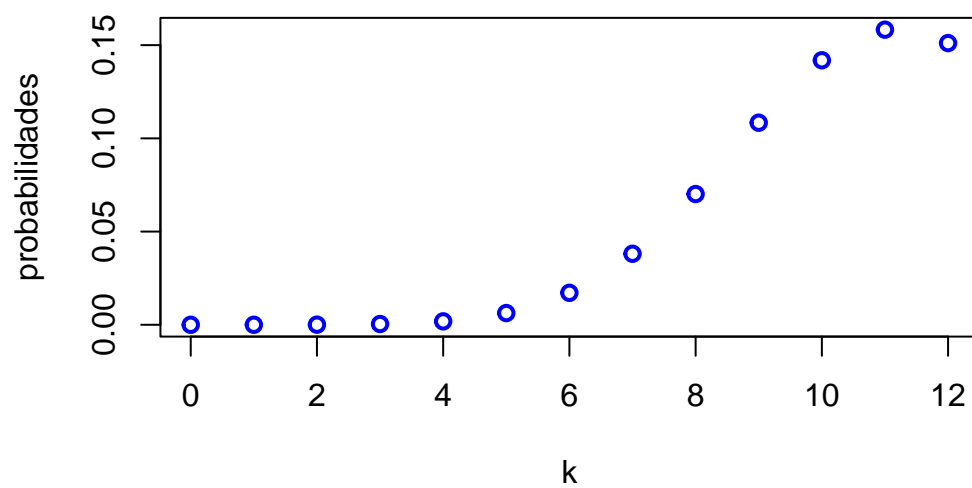
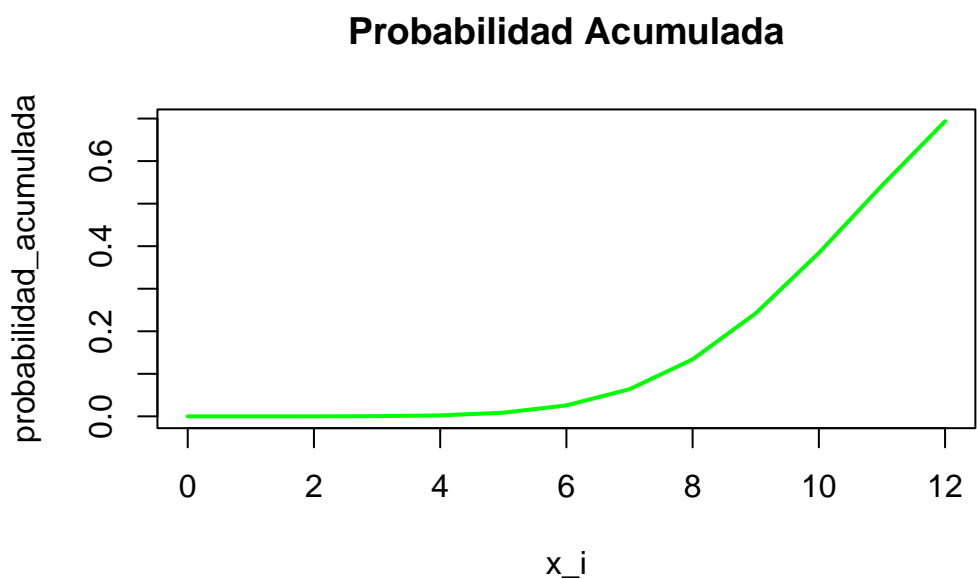


Gráfico De Probabilidades





garrapatas

```
[[1]]
```

	x	Coefficiente	Probabilidad	x_i	coeficiente_i	Probabilidades	Acumulada
1	12	5200300	0.1511101	0	1	3.228994e-07	3.228994e-07
2	12	5200300	0.1511101	1	25	6.604760e-06	6.927660e-06
3	12	5200300	0.1511101	2	300	6.484674e-05	7.177440e-05
4	12	5200300	0.1511101	3	2300	4.067659e-04	4.785403e-04
5	12	5200300	0.1511101	4	12650	1.830447e-03	2.308987e-03
6	12	5200300	0.1511101	5	53130	6.290080e-03	8.599067e-03
7	12	5200300	0.1511101	6	177100	1.715476e-02	2.575383e-02
8	12	5200300	0.1511101	7	480700	3.809694e-02	6.385077e-02
9	12	5200300	0.1511101	8	1081575	7.013301e-02	1.339838e-01
10	12	5200300	0.1511101	9	2042975	1.083874e-01	2.423712e-01
11	12	5200300	0.1511101	10	3268760	1.418889e-01	3.842601e-01
12	12	5200300	0.1511101	11	4457400	1.583058e-01	5.425659e-01
13	12	5200300	0.1511101	12	5200300	1.511101e-01	6.936760e-01

```
[[2]]
$breaks
[1] 0.0 0.2 0.4 0.6 0.8 1.0
```

```

$counts
[1] 13  0  0  0 12

$density
[1] 2.6 0.0 0.0 0.0 2.4

$mids
[1] 0.1 0.3 0.5 0.7 0.9

$xname
[1] "y"

$equidist
[1] TRUE

attr(,"class")
[1] "histogram"

[[3]]
NULL

```

Conclusiones

La utilidad de la distribución binomial se extiende a diferentes campos de estudio y aplicaciones prácticas. Nos permite realizar inferencias estadísticas, evaluar la eficacia de tratamientos médicos en ensayos clínicos, estudiar la probabilidad de eventos binarios en genética, controlar la calidad en la industria, valorar opciones financieras y mucho más.

Utilizar la distribución binomial nos proporciona una herramienta poderosa para comprender, modelar y analizar eventos y procesos que involucran resultados binarios. Su aplicación nos permite tomar decisiones informadas, realizar predicciones y realizar análisis de datos que son fundamentales en una amplia gama de disciplinas y contextos profesionales.

Literatura citada

- Altham, Patricia ME. 2018. “Two Generalizations of the Binomial Distribution.” *Journal of the Royal Statistical Society Series C: Applied Statistics* 27 (2): 162–67.
- Georgacanos. 1988. *Probabilidad y Estadística*. Vol. 1. McGRA W-Hill.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.