

PRINCIPALES DISTRIBUCIONES DISCRETAS

Objetivos generales del tema

En este tema definiremos y discutiremos diversas e importantes distribuciones discretas, es decir, funciones masa de probabilidad o funciones de distribución correspondientes a variables aleatorias discretas. Veremos las propiedades más importantes.

Principales contenidos

- Distribución de Bernoulli.
- Distribución Binomial.
- Distribución Binomial Negativa.
- Distribución Hipergeométrica.
- Distribución de Poisson.

1 Introducción

Recordemos que una variable aleatoria X tiene una distribución de probabilidad discreta (o simplemente una distribución discreta) si:

$$P(X = x_i) = \begin{cases} p_i & \text{si } x = x_i \text{ (} i = 1, 2, \dots \text{)} \\ 0 & \text{en el otro caso} \end{cases} \quad \text{con } p_i \geq 0 \text{ y } \sum_i p_i = 1$$

2 Distribución de Bernoulli

La distribución de Bernoulli es una distribución que toma dos valores (1 y 0) con probabilidad p y $1-p$. Habitualmente esta distribución aparece en experimentos donde se identifica el suceso 1=EXITO y 0=FRACASO.

Definición 1 Una v.a. X tiene una distribución de Bernoulli si (para algún p con $0 \leq p \leq 1$)

$$\begin{aligned} P(X = 1) &= p \\ P(X = 0) &= 1 - p \end{aligned}$$

Ejemplo 2 La v.a. que define el experimento lanzamiento de una moneda sigue una distribución de Bernoulli de parámetro p . Donde p es la probabilidad del suceso de interés, cara o cruz.

PROPIEDADES

<i>Distribución de Bernoulli de parámetro p</i>	
Media μ	$\mathbb{E}(X) = p$
Varianza σ^2	$\text{Var}(X) = p(1-p)$
Desviación Típica σ	$\sqrt{p(1-p)}$

3 Distribución Binomial

La distribución Binomial de parámetro n y p ($Bi(n, p)$) surge como una secuencia n intentos del tipo de Bernoulli que verifica:

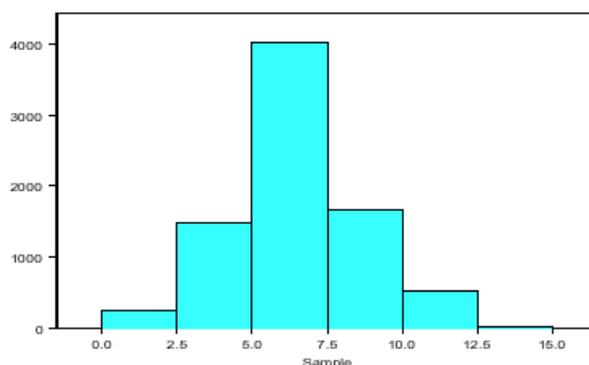
- Los intentos son independientes.
- Cada resultado del intento puede tomar únicamente dos resultados mutuamente excluyentes, que denotaremos por EXITO (E) o FRACASO (F).
- La probabilidad de éxito (y por lo tanto la de fracaso) es constante en cada intento.

Así cada intento, definiendo $p = P(\text{EXITO})$ y asignando el valor 0 cuando ocurre F y el valor 1 cuando ocurre E, se puede describir como una distribución de Bernoulli de parámetro p .

Definición 3 Una v.a. X tiene una distribución Binomial si (para algún entero positivo n y algún p con $0 \leq p \leq 1$) mide

$X =$ "nº de éxitos de un total de n si la probabilidad de éxito es p "

$$P(X = k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{si } k = 1, 2, \dots, n \\ 0 & \text{en el otro caso} \end{cases}$$



Distribución Binomial $n = 25$, $p = 0.25$.

La forma de la distribución binomial va aumentando hasta $m = (n + 1)p$ y después comienza a bajar.

Una v.a. binomial puede ser considerada como una suma de n variables aleatorias de Bernoulli de parámetro p , es decir, como la variable indicadora del número de éxitos en n pruebas de Bernoulli, es decir, si denotamos por X_i el i -ésimo intento de Bernoulli, entonces, $X = \sum_{i=1}^n X_i$.

Observación 4 La distribución de Bernoulli es un caso particular de la distribución Binomial con $n = 1$.

Proposición 5 Si X tiene la distribución Binomial con n (número de intentos) y p (probabilidad de éxito que denotaremos por $X \sim Bi(n, p)$) entonces la variable $Y = n - X$, que representa el número de fracasos, es una variable de Binomial con n (número de intentos) y probabilidad de éxito $1 - p$.

$$X \sim Bi(n, p) \Rightarrow Y = n - X \sim Bi(n, 1 - p)$$

PROPIEDADES

<i>Distribución Binomial de parámetros n y p</i>	
Media μ	$\mathbb{E}(X) = np$
Varianza σ^2	$Var(X) = np(1 - p)$
Desviación Típica σ	$\sqrt{np(1 - p)}$

Ejemplo 6 *En una fábrica hay 12 máquinas. Cada una de ellas está averiada un día de cada 10. ¿Cuál es la probabilidad de que un determinado día haya más de 3 máquinas averiadas?*

Sea la variable $X = \text{"nº de máquinas averiadas de un total de 12 si } P(\text{avería}) = 0,1 \text{"} \in Bi(n = 12, p = 0.1)$.

$$\begin{aligned} P(X > 3) &= 1 - P(X \leq 3) = 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)] = \\ &= 1 - \left[\binom{12}{0} 0.1^0 (1 - 0.1)^{12-0} + \binom{12}{1} 0.1^1 (1 - 0.1)^{12-1} + \right. \\ &\quad \left. + \binom{12}{2} 0.1^2 (1 - 0.1)^{12-2} + \binom{12}{3} 0.1^3 (1 - 0.1)^{12-3} \right] = 1 - 0.9744 = 0.0256. \end{aligned}$$

La v.a. binomial está tabulada para valores pequeños de n . Hay tablas que dan la función de masa de probabilidad $P(X = k)$ o bien la función de distribución $P(X \leq k)$, aunque sólo vienen para valores $p \leq 0.5$.

4 Distribución Binomial Negativa

La distribución Binomial Negativa de parámetro r y p ($BN(r, p)$) surge como una secuencia infinita de intentos del tipo de Bernoulli que verifican:

- La secuencia de intentos es independiente.
- Cada resultado del intento puede tomar únicamente dos resultados mutuamente excluyentes, que denotaremos por EXITO (E) o FRACASO (F).
- La probabilidad de éxito (y por lo tanto la de fracaso) es constante en cada intento.
- Los intentos continúan (se ejecutan) hasta que un total de r éxitos se hayan observado.

Así cada intento, definiendo $p = P(\text{EXITO})$ y asignando el valor 0 cuando ocurre F y el valor 1 cuando ocurre E, se puede describir como una distribución de Bernoulli de parámetro p .

Definición 7 *Una v.a. X tiene una distribución Binomial Negativa si (para algún entero positivo r y algún p con $0 \leq p \leq 1$)*

$X = \text{"nº de fracasos hasta el } r - \text{ésimo éxito si } p = P(\text{EXITO}) \text{"}$

$$P(X = k) = \begin{cases} \binom{r+k-1}{k} p^r (1-p)^k & \text{si } k = 0, 1, 2, \dots \\ 0 & \text{en el otro caso} \end{cases}$$

Proposición 8 Si X tiene la distribución Binomial Negativa con r (número de éxitos) y p (probabilidad de éxito) que denotaremos por $X \sim BN(r, p)$ entonces la variable $Y = X + r$, que representa el número de pruebas para del éxito r -ésimo, tiene una distribución relacionada con X por:

$$P(Y = k) = P(X = k - r) = \begin{cases} \binom{k-1}{k-r} p^r (1-p)^{k-r} & \text{si } k = r, r+1, r+2, \dots \\ 0 & \text{en el otro caso} \end{cases}$$

Observación 9 La distribución Binomial Negativa con $r = 1$ es conocida en la literatura por **distribución geométrica** de parámetro p y mide el número de fracasos antes del primer éxito.

PROPIEDADES

Distribución Binomial Negativa de parámetros r y p

Media μ	$\mathbb{E}(X) = r \frac{(1-p)}{p}$
Varianza σ^2	$Var(X) = r \frac{(1-p)}{p^2}$
Desviación Típica σ	$\frac{\sqrt{r(1-p)}}{p}$

5 Distribución Hipergeométrica

La distribución Hipergeométrica de parámetros k , n y N ($HG(k, n, N)$) surge en situaciones en donde el modelo aproximado de probabilidad se corresponde con muestreo sin reemplazamiento de una población dicotómica (Éxito y Fracaso) finita. Concretamente, las suposiciones que llevan a considerar esta distribución son:

- La población o conjunto donde deba hacerse el muestreo consta de N individuos o elementos a seleccionar.
- Cada individuo puede ser caracterizado como un éxito (E) o fracaso (F).
- Se selecciona una muestra de n individuos de entre los k individuos marcados como éxito y los $N - k$ restantes marcados como fracaso.
- Hay selección equiprobable en cada paso.

Definición 10 Una v.a. X tiene una distribución Hipergeométrica si (para algunos enteros positivos k , n y N)

$$X = \text{"n}^\circ \text{ de individuos de un total de } n \text{ con cierta característica (éxito) si en } N \text{ hay un total de } k \text{"}$$

$$P(X = x) = \begin{cases} \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} & \text{si } \max\{0, n - (n - k)\} \leq x \leq \min(n, k) \\ 0 & \text{en el otro caso} \end{cases}$$

Proposición 11 Podemos aproximar la distribución Hipergeométrica a la distribución Binomial, puesto que cuando N es grande ($N > 10n$) podemos suponer que $X \in Bi(n, p)$ con $p = \frac{k}{N}$.

PROPIEDADES

Distribución Hipergeométrica de parámetros k, n y N

Media μ	$\mathbb{E}(X) = np$
Varianza σ^2	$Var(X) = npq \frac{N-n}{N-1}$
Desviación Típica σ	$\sqrt{npq \frac{N-n}{N-1}}$

Ejemplo 12 De cada 2.000 tornillos fabricados por una determinada máquina hay 2 defectuosos. Para realizar el control de calidad se observan 150 tornillos y se rechaza el lote si el número de defectuosos es mayor que 1. Calcular la probabilidad de que el lote sea rechazado.

Sea $X =$ "nº de defectuosos de un total de $n = 150$ si en $N = 2.000$ hay un total de $k = 2 \in HG(k = 2, n = 150, N = 2.000)$

Se pide $P(X > 1) = 1 - [P(X = 0) + P(X = 1)] =$

$$= 1 - \left[\frac{\binom{2}{0} \binom{2.000-2}{150}}{\binom{2.000}{150}} + \frac{\binom{2}{1} \binom{2.000-2}{150-1}}{\binom{2.000}{150}} \right]$$

Como los cálculos son complicados, se va a aproximar la v.a. X por una binomial $Bi(n = 150, p = \frac{2}{2.000} = 0.001)$

$$P(X > 1) = 1 - [P(X = 0) + P(X = 1)] = 1 - \left[\binom{150}{0} 0.001^0 (1 - 0.001)^{150-0} + \binom{150}{1} 0.001^1 (1 - 0.001)^{150-1} \right] \approx 0.442$$

6 Distribución de Poisson

Consideremos sucesos ("éxitos") que ocurren aleatoriamente a lo largo del tiempo y sea la v.a. $X =$ "nº de sucesos en el intervalo de tiempo $(0, t)$ ". La distribución de Poisson de parámetro λ surge en los llamados procesos de Poisson que se presenta en relación con el acontecimiento de éxitos de un tipo particular

durante un intervalo continuo de tiempo. Las suposiciones de un proceso de Poisson, durante un intervalo de tiempo que empieza en $t = 0$ son:

- Existe un parámetro λ tal que para cualquier intervalo corto de tiempo Δt , la probabilidad de que ocurra exactamente un éxito es $\lambda \Delta t$.
- La probabilidad de que se produzca más de un éxito en el intervalo Δt es $o(\Delta t)^1$.
- El número de éxitos ocurrido en un intervalo Δt es independiente del número ocurrido antes de este intervalo de tiempo.

Definición 14 Una v.a. X tiene una distribución de Poisson si (para algún λ con $\lambda \geq 0$)

$X =$ "nº de éxitos en un intervalo de tiempo si el nº medio es λ "

$$P(X = k) = \begin{cases} e^{-\lambda} \frac{\lambda^k}{k!} & \text{si } k = 0, 1, 2, \dots \\ 0 & \text{en el otro caso} \end{cases}$$

Proposición 15 Sea $X \sim \text{Poisson}(\lambda_1)$, $Y \sim \text{Poisson}(\lambda_2)$ independientes, entonces la variable $X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$

Proposición 16 Se puede considerar que una v.a. $\text{Poisson}(\lambda)$ es límite de una binomial $Bi(n, p)$ si

$$p < 0.1 \quad \text{y} \quad np < 5$$

Proposición 17 La distribución de Poisson presenta su máximo en los puntos $x = \lambda$ y $x = \lambda - 1$.

1

Definición 13 $o(\Delta t)$ si $\lim_{t \rightarrow \infty} \frac{o(\Delta t)}{\Delta t} = 0$

PROPIEDADES

Distribución de Poisson de parámetros λ

Media μ	$\mathbb{E}(X) = \lambda$
Varianza σ^2	$\text{Var}(X) = \lambda$
Desviación Típica σ	$\sqrt{\lambda}$

Ejemplo 18 *En un taller se averían una media de 2 máquinas a la semana. Calcula la probabilidad de que no haya ninguna avería en una semana. ¿Y de que haya menos de 6 en un mes?*

Sea $X =$ "nº de averías en una semana si de media hay $\lambda = 2 \in \text{Poisson}(\lambda = 2)$.

Entonces

$$P(X = 0) = e^{-2} \frac{2^0}{0!} = 0.14$$

Si en una semana hay 2 averías de media, entonces en un mes hay $2 \times 4 = 8$ averías de media. Sea entonces $Y =$ "nº de averías en un mes si de media hay $\lambda = 8 \in \text{Poisson}(\lambda = 8)$

$$P(Y \leq 5) = P(Y = 0) + \dots + P(Y = 5) = e^{-2} \frac{2^0}{0!} + \dots + e^{-2} \frac{2^5}{5!} = 0.191$$