Tema 6: MODELOS DE PROBABILIDAD DISCRETOS

En este tema se van a estudiar algunas distribuciones de probabilidad de tipo discreto que van a permitir modelizar gran número de situaciones reales. En particular estudiaremos su función masa de probabilidad así como su función de distribución y algunas de sus características, como la esperanza, la varianza y la función generatriz de momentos. Todos estos modelos dependerán de algunas constantes que reciben el nombre de parámetros de la distribución.

1. Distribución degenerada
2. Distribución uniforme discreta.
3. Distribución de Bernouilli.
4. Distribución binomial.
5. Distribución de Poisson.
6. Distribución binomial negativa.
7. Distribución hipergeométrica.

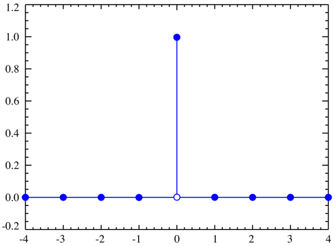
1. Distribución degenerada

Se aplica este modelo a cualquier situación en la que se tiene un experimento determinístico que produce siempre el mismo resultado.

Se define la variable aleatoria discreta X, que representa el resultado del experimento, como la variable que toma un único valor, c, con probabilidad 1.

Definición

Sea X una variable aleatoria que sólo toma el valor c. Se dice que X tiene una distribución degenerada en c si $P[X = x] = 1, \quad x = c.$



- Función de distribución: $F(x) = P[X \leq x] = \begin{cases} 0, & \text{si } x < c \\ 1, & \text{si } x \geq c \end{cases}$
- Media y varianza: $E[X] = c$ y $Var[X] = 0$.
- Función generatriz de momentos: $M_X(t) = E[e^{tX}] = e^{tc}, \quad t \in \mathbb{R}$.
- Momentos: $m_k = E[X^k] = c^k$ y $\mu_k = E[(X-c)^k] = 0$.
- Caracterización: X es degenerada si y sólo si $Var[X] = 0$.

Ejemplo

Se considera el experimento de lanzar una moneda con doble cara. Determinar la variable aleatoria asociada al experimento así como sus características más importantes.

2. Distribución uniforme discreta

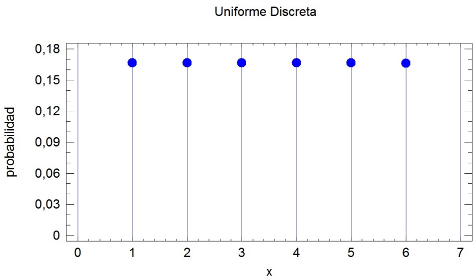
Se aplica este modelo a cualquier situación en la que se tiene un experimento cuyos posibles resultados sea un conjunto finito formado por sucesos equiprobables.

Se define la variable aleatoria discreta X, que representa el resultado del experimento, como la variable que toma un número finito de valores, $\{x_1, \dots, x_n\}$, todos con la misma probabilidad.

Definición

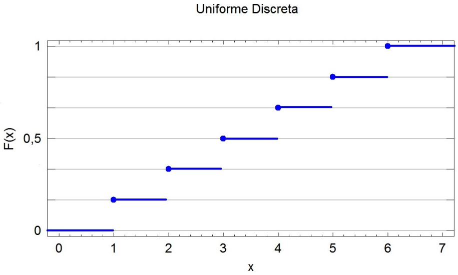
Sea X una variable aleatoria que toma valores en $\{x_1, \ldots, x_n\}$, verificando $x_i \in \mathbb{R}$, $i = 1, \ldots, n$, $x_1 < x_2 < \cdots < x_n$. Se dice que X tiene una distribución uniforme discreta en dichos puntos y se denota como $X \rightsquigarrow U(x_1, \ldots, x_n)$ si

$$P[X = x_i] = \frac{1}{n}, \; i = 1, \ldots, n.$$



- Función de distribución:

$$F(x) = P[X \leq x] = \begin{cases} 0, & \text{si } x < x_1 \\ \dfrac{i-1}{n}, & \text{si } x_{i-1} \leq x < x_i, \; i = 2, \ldots, n \\ 1, & \text{si } x \geq x_n \end{cases}$$



- Media y varianza:

$$E[X] = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x} \quad \text{y} \quad Var[X] = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sigma^2.$$

- Función generatriz de momentos:

$$M_X(t) = E[e^{tX}] = \frac{1}{n}\sum_{i=1}^{n} e^{tx_i}, \quad t \in \mathbb{R}.$$

- Momentos:

$$m_k = E[X^k] = \frac{1}{n}\sum_{i=1}^{n} x_i^k \quad \text{y} \quad \mu_k = E[(X - \bar{x})^k] = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^k.$$

Ejemplo

Se considera el experimento de lanzar un dado. Determinar la variable aleatoria asociada al experimento así como sus características más importantes.

3. Distribución de Bernouilli

Se aplica este modelo a cualquier situación en la que se tiene un experimento que produce sólo dos resultados posibles e incompatibles, a los que se denominan **éxito** y **fracaso**.

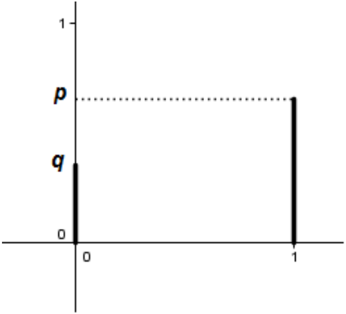
Se define la variable aleatoria discreta X, que representa el resultado del experimento, como la variable que asigna un 1 al suceso éxito con probabilidad p y un 0 al suceso fracaso con probabilidad $q = 1 - p$.

$$X : \begin{cases} 0, & \text{si ocurre el fracaso, con probabilidad } q \\ 1, & \text{si ocurre el éxito, con probabilidad } p \end{cases}$$

Definición

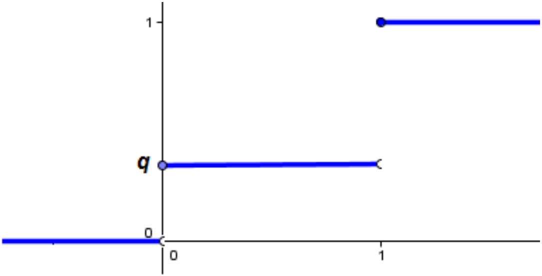
Sea X *una variable aleatoria que toma los valores 0 ó 1. Se dice que* X *tiene una distribución de* **Bernouilli** *de parámetro* p, *y se denota como* $X \rightsquigarrow B(1,p)$, $(0 < p < 1)$ *si*

$$P[X = x] = p^x(1-p)^{1-x} \qquad \forall x = 0, 1$$



- Función de distribución

$$F(x) = P[X \leq x] = \begin{cases} 0, & \text{si } x < 0 \\ 1-p, & \text{si } 0 \leq x < 1 \\ 1, & \text{si } x \geq 1 \end{cases}$$



- Esperanza y varianza: $E[X] = p$ y $Var[X] = p(1-p)$.
- Función generatriz de momentos:

$$M_X(t) = E[e^{tX}] = pe^t + (1-p), \quad t \in \mathbb{R}.$$

- Momentos:

$$m_k = E[X^k] = p \quad \text{y} \quad \mu_k = E[(X-p)^k] = (1-p)^k p + (-p)^k (1-p).$$

Ejemplo

Se considera el experimento de lanzar una moneda y comprobar si sale cara. Determinar la variable aleatoria asociada al experimento así como sus características más importantes.

4. Distribución binomial

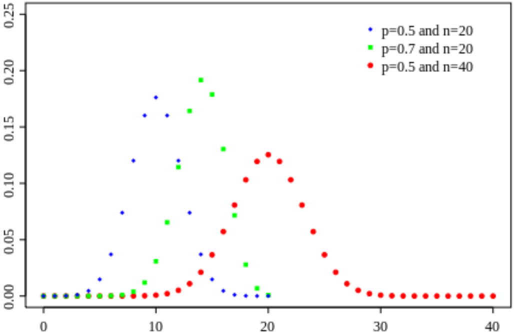
Se aplica este modelo a cualquier situación donde se considere un experimento que produce sólo dos resutlados posibles e incompatibles y que se repite n veces de forma idéntica e independiente. Los resultados de cada realización del experimento se clasifican en dos categorías, éxito y fracaso, como en el caso de la distribución de Bernoulli.

Se define la variable aleatoria discreta X como aquella que cuenta el número de éxitos en las n realizaciones del experimento, siendo p la probabilidad de éxito del experimento.

Definición

Sea X una variable aleatoria que toma valores en $\{0, 1, 2, 3, \ldots, n\}$. Se dice que X tiene una distribución **binomial** *de parámetros n y p, y se denota como $X \rightsquigarrow B(n, p)$, $n \in \mathbb{N}$ y $(0 < p < 1)$ si*

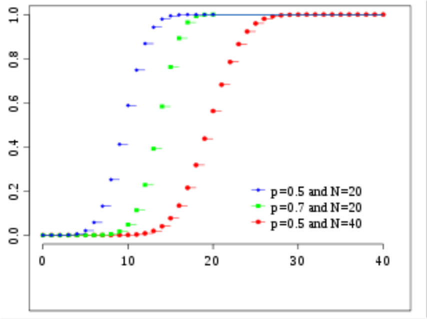
$$P[X = x] = \binom{n}{x} p^x (1-p)^{n-x} \qquad \forall x = 0, \ldots, n.$$



- Función de distribución

$$F(x) = P[X \leq x] = \begin{cases} 0, & x < 0 \\ \displaystyle\sum_{k=0}^{[x]} \binom{n}{k} p^k (1-p)^{n-k}, & 0 \leq x < n \\ 1, & x \geq n \end{cases}$$

donde $[x]$ denota la parte entera de x.



- Esperanza y varianza: $E[X] = np$ y $Var[X] = np(1-p)$.
- Función generatriz de momentos:

$$M_X(t) = E[e^{tX}] = (pe^t + (1-p))^n, \quad t \in \mathbb{R}.$$

- Propiedad de simetría: $X \rightsquigarrow B(n,p) \Leftrightarrow Y = n - X \rightsquigarrow B(n, 1-p)$.
- La función masa de probabilidad está tabulada para distintos valores de los parámetros (ver tablas).

Ejemplo

Se repite el experimento anterior 10 veces usando la misma moneda. Determinar la variable aleatoria asociada al experimento y su función masa de probabilidad. Calcular:

a) *la probabilidad de que se obtengan exactamente 5 caras,*

b) *el número esperado de caras que se obtendrán y su varianza,*

c) *la probabilidad de que se obtenga a lo sumo 1 cara,*

d) *la probabilidad de que se obtenga al menos una cara.*

Ejemplo

Ahora se cambia la moneda a una cargada cuya probabilidad de salir cara es 0.7 y se realizan 12 lanzamientos. Calcular:

a) *la probabilidad de que se obtengan exactamente 5 caras,*

b) *la probabilidad de que se obtengan más de 7 caras.*

5. Distribución de Poisson

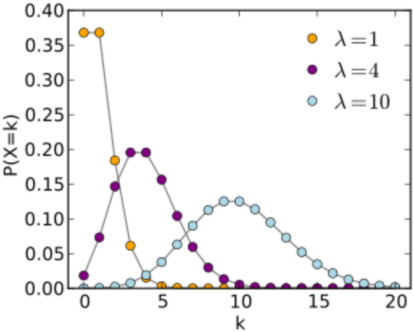
Se aplica este modelo a cualquier situación donde se considere un suceso que ocurre de manera aleatoria en el tiempo o en el espacio con una probabilidad pequeña de ocurrencia (ley de los sucesos raros).

Se define la variable aleatoria discreta X como aquella que cuenta el número de veces que se presenta el suceso en un intervalo de tiempo dado o región del espacio fijada siendo λ el número medio de ocurrencias del suceso considerado en el intervalo de tiempo o espacio fijado.

Definición

Sea X una variable aleatoria discreta que toma los valores 0, 1, 2, ... Se dice que X tiene una distribución de Poisson de parámetro λ, y se denota como $X \rightsquigarrow \mathcal{P}(\lambda)$, $(\lambda > 0)$ si

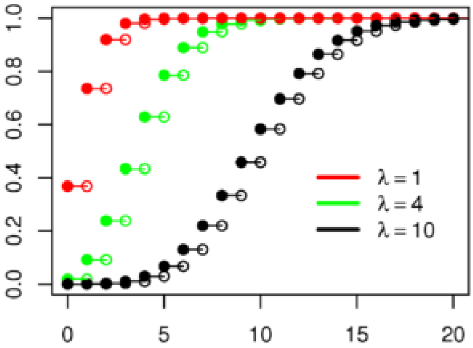
$$P[X = k] = e^{-\lambda}\frac{\lambda^k}{k!} \qquad \forall k = 0, 1, 2, \ldots$$



- Función de distribución

$$F(x) = P[X \leq x] = \begin{cases} 0, & x < 0 \\ \displaystyle\sum_{k=0}^{[x]} \frac{\lambda^k}{k!} e^{-\lambda}, & x \geq 0 \end{cases}$$

donde $[x]$ es la parte entera de x.



- Esperanza y varianza: $E[X] = \lambda$ y $Var[X] = \lambda$.
- Función generatriz de momentos:

$$M_X(t) = E[e^{tX}] = exp(\lambda(e^t - 1)), \quad t \in \mathbb{R}.$$

- La función masa de probabilidad está tabulada para distintos valores del parámetro (ver tablas).

Ejemplo

En una asignatura los alumnos tienen la posibilidad de salir voluntariamente a la pizarra en las clases prácticas para ser evaluados en la componente de participación. En particular se sabe que, en media, un alumno sale 3 veces al curso en dicha asignatura. Determinar la variable aleatoria asociada, así como la función masa de probabilidad. Calcular:

a) *el número esperado de veces que un alumno sale a la pizarra en un curso y su varianza,*
b) *la probabilidad de que un alumno salga exactamente tres veces en un curso,*
c) *la probabilidad de que un alumno salga más de 5 veces en un curso.*

Aproximación a la distribución de Poisson

La distribución de Poisson surgió como aproximación de la distribución $B(n, p)$ ya que su f.m.p. es el límite (cuando $n \to \infty$ y $p \to 0$) de la correspondiente para $\lambda = n \times p$.
Por ello, si se tiene $X \rightsquigarrow B(n, p)$, con n suficientemente grande y p suficientemente pequeño, entonces se puede aproximar esta distribución a una distribución de Poisson para el cálculo de las probabilidades asociadas, es decir,

$$X \rightsquigarrow B(n, p) \Longrightarrow P[X = x] \approx P[Y = x] \text{ donde } Y \rightsquigarrow \mathcal{P}(\lambda = n \times p).$$

En la práctica se ha comprobado que se puede aceptar que dicha aproximación se verifica si:

- $n \geq 30$,
- $p \leq 0.1$ (ó $1 - p \leq 0.1$),
- $n \times p < 5$.

Ejemplo

De nuevo se lanza un moneda pero está cargada de forma que la probabilidad de salir cara es sólo de 0.1. Calcular la probabilidad de que al lanzar la moneda 40 veces se obtengan 5 caras.

6. Distribución binomial negativa

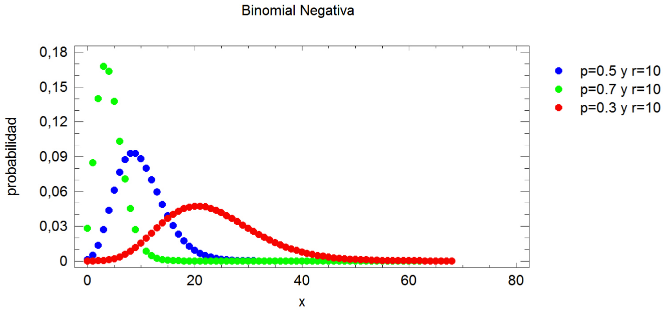
Se aplica este modelo a cualquier situación donde se considere un experimento que produce sólo dos resultados posibles e incompatibles y que se repite, de forma idéntica e independiente, hasta que ocurre un número especificado de éxitos. Los resultados de cada realización del experimento se clasifican en dos categorías, éxito y fracaso, como en el caso de la distribución de Bernoulli.

Se define la variable aleatoria discreta X como aquella que cuenta el número de fracasos hasta que ocurren r éxitos, siendo p la probabilidad de éxito del experimento.

Definición

Sea X *una variable aleatoria que toma valores en* $\{0, 1, 2, \ldots\}$. *Se dice que* X *tiene una distribución* **binomial negativa** *de parámetros* r *y* p, *y se denota como* $X \rightsquigarrow BN(r, p)$, $r \in \mathbb{N}$ *y* $(0 < p < 1)$ *si*

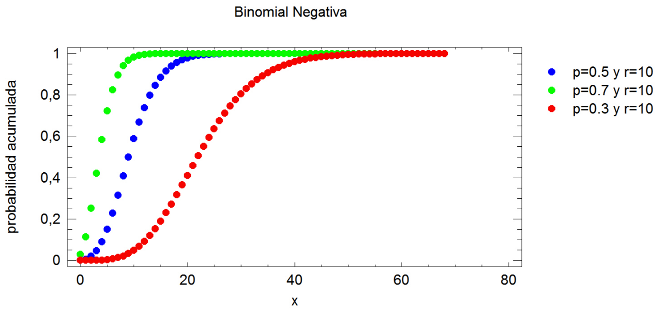
$$P[X = x] = \binom{x + r - 1}{x} (1 - p)^x p^r \qquad \forall x = 0, 1, 2 \ldots .$$



- Función de distribución

$$F(x) = P[X \leq x] = \begin{cases} 0, & x < 0 \\ p^r \displaystyle\sum_{k=0}^{[x]} \binom{k+r-1}{k} (1-p)^k, & x \geq 0 \end{cases}$$

donde $[x]$ denota la parte entera de x.



- Esperanza y varianza: $E[X] = \dfrac{r(1-p)}{p}$ y $Var[X] = \dfrac{r(1-p)}{p^2}$.
- Función generatriz de momentos:

$$M_X(t) = E[e^{tX}] = p^r(1-(1-p)e^t)^{-r}, \quad t < -\ln(1-p).$$

- Relación con la binomial:

$X \rightsquigarrow BN(r,p) \Rightarrow P[X=x] = P[Y=r-1]p$ con $Y \rightsquigarrow B(r+x-1,p)$.

- Caso particular: la distribución geométrica (r=1).

Ejemplo

Se repite el experimento de lanzar una moneda normal hasta obtener 5 caras. Determinar la variable aleatoria asociada al experimento y su función masa de probabilidad. Calcular:

a) *la probabilidad de que se obtengan exactamente 5 cruces,*
b) *el número esperado de cruces que se obtendrán y su varianza,*
c) *la probabilidad de que se obtenga a lo sumo una cruz,*
d) *la probabilidad de que se obtenga al menos una cruz.*

6.1. Distribución geométrica

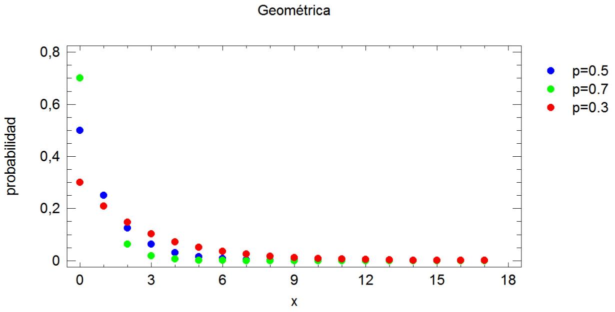
Se aplica este modelo a cualquier situación donde se considere un experimento que produce sólo dos resultados posibles e incompatibles y que se repite, de forma idéntica e independiente, hasta que ocurre el primer éxito. Los resultados de cada realización del experimento se clasifican en dos categorías, éxito y fracaso, como en el caso de la distribución de Bernoulli.

Se define la variable aleatoria discreta X como aquella que cuenta el número de fracasos hasta que ocurren el primer éxito, siendo p la probabilidad de éxito del experimento.

Definición

Sea X *una variable aleatoria que toma valores en* $\{0, 1, 2, \ldots\}$. *Se dice que* X *tiene una distribución* **geométrica** *de parámetro* p, *y se denota como* $X \rightsquigarrow G(p), (0 < p < 1)$ *si*

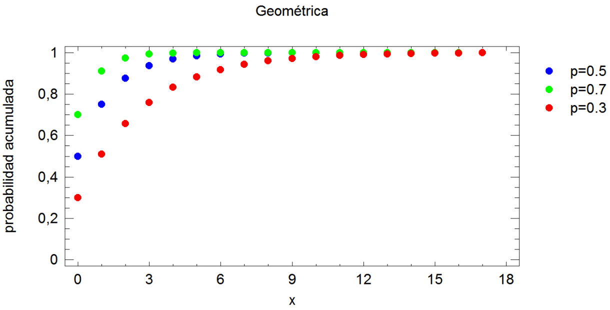
$$P[X = x] = (1-p)^x p \qquad \forall x = 0, 1, 2 \ldots .$$



- Función de distribución

$$F(x) = P[X \leq x] = \begin{cases} 0, & x < 0 \\ 1 - (1-p)^{[x]+1}, & x \geq 0 \end{cases}$$

donde $[x]$ denota la parte entera de x.



- Esperanza y varianza: $E[X] = \dfrac{(1-p)}{p}$ y $Var[X] = \dfrac{(1-p)}{p^2}$.
- Función generatriz de momentos:

$$M_X(t) = E[e^{tX}] = p(1-(1-p)e^t)^{-1}, \quad t < -\ln(1-p).$$

- Falta de memoria:

$$X \rightsquigarrow G(p) \Rightarrow P[X \geq h+k/X \geq h] = P[X \geq k], \quad \forall h, k \in \mathbb{N} \cup \{0\}.$$

- Caracterización: La distribución geométrica es la única con valores en $\mathbb{N} \cup \{0\}$ que verifica la propiedad de falta de memoria.

Ejemplo

Se repite el experimento de lanzar una moneda cargada con probabilidad de cara de 0.1 hasta obtener la cara. Determinar la variable aleatoria asociada al experimento y su función masa de probabilidad. Calcular:

a) *la probabilidad de que se obtengan exactamente 5 cruces,*
b) *el número esperado de cruces que se obtendrán y su varianza,*
c) *la probabilidad de que se obtengan a lo sumo 10 cruces,*
d) *la probabilidad de que se obtengan al menos 10 cruces sabiendo que en los primeros cinco lanzamientos no se obtuvo ninguna cara.*

7. Distribución hipergeométrica

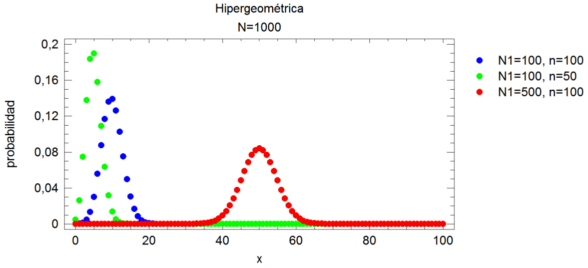
Se aplica este modelo para el muestreo aleatorio sin reemplazamiento en poblaciones finitas. En particular modeliza el número de individuos en una muestra que presentan una determinada característica.

Se considera una población de tamaño N y una subpoblación de esta de tamaño N_1. Se toma una muestra de tamaño n de la población sin reemplazamiento o simultáneamente. Entonces se define la variable aleatoria discreta X como aquella que cuenta el número de individuos de la muestra que pertenecen a la subpoblación.

Definición

Sea X *una variable aleatoria discreta que toma los valores en* $\{0, 1, 2, \ldots, n\}$. *Se dice que* X *tiene una distribución hipergeométrica de parámetros* N, N_1 *y* n, *y se denota como* $X \rightsquigarrow H(N, N_1, n)$, $N, N_1, n \in \mathbb{N}$, $n, N_1 \leq N$ *si*

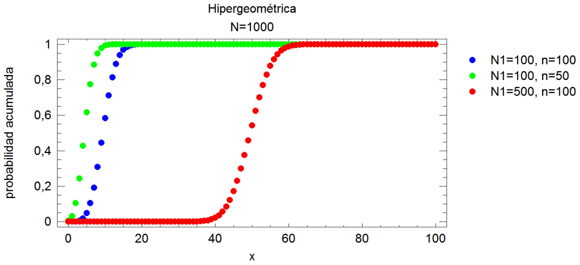
$$P[X = x] = \frac{\binom{N_1}{x}\binom{N-N_1}{n-x}}{\binom{N}{n}} \qquad \max(0, n-(N-N_1)) \leq x \leq \min(n, N_1)$$



- Función de distribución

$$F(x) = P[X \leq x] = \begin{cases} 0, & x < \max(0, n-(N-N_1)) \\ \displaystyle\sum_{k=0}^{[x]} \frac{\binom{N_1}{k}\binom{N-N_1}{n-k}}{\binom{N}{n}}, & \max(0, n-(N-N_1)) \leq x < \min(n, N_1) \\ 1, & x \geq \min(n, N_1) \end{cases}$$

donde $[x]$ es la parte entera de x.



- Esperanza y varianza: $E[X] = \frac{nN_1}{N}$ y $Var[X] = \frac{nN_1}{N}\left(1 - \frac{N_1}{N}\right)\frac{N-n}{N-1}$.
- Función generatriz de momentos:

$$M_X(t) = E[e^{tX}] = \sum_{x=\text{máx}(0,n-(N-N_1))}^{\text{mín}(n,N_1)} e^{tx} \frac{\binom{N_1}{x}\binom{N-N_1}{n-x}}{\binom{N}{n}}, \quad t \in \mathbb{R}.$$

Ejemplo

En una clase de 100 alumnos 20 de ellos son repetidores. Se seleccionan 10 alumnos de la clase y se les pregunta si están repitiendo. Determinar la variable aleatoria asociada, así como la función masa de probabilidad. Calcular:

a) *el número esperado de repetidores entre los seleccionados y su varianza,*
b) *la probabilidad de que tres de los seleccionados sean repetidores.*

Relación con la distribución binomial

Existe una relación entre la distribución hipergeométrica y la distribución binomial ya que el límite (cuando $N \to \infty$) de la f.m.p. de la hipergemétrica resulta ser la f.m.p. de un binomial con parámetros n y $p = N_1/N$. Por ello, si se tiene $X \rightsquigarrow H(N, N_1, n)$, con N suficientemente grande, entonces se puede aproximar esta distribución a una distribución binomial para el cálculo de las probabilidades asociadas, es decir,

$$X \rightsquigarrow H(N, N_1, n) \Longrightarrow P[X = x] \approx P[Y = x] \text{ donde } Y \rightsquigarrow B(n, N_1/N).$$

En la práctica se ha comprobado que se puede aceptar que dicha aproximación se verifica si:

- $N > 50$,
- $n \leq 0.1N$.

Ejemplo

Repetir el ejemplo anterior usando la aproximación para los cálculos.