TEMA 6 - VARIABLE DISCRETA

Experimento de Bernouilli

Un experimento de Bernouilli es un experimento aleatorio que da lugar únicamente a dos sucesos A y $\bar{A}$ que son mutuamente excluyentes ($A \cap \bar{A} = \emptyset$) y exhaustivos ($A \cup \bar{A} = \Omega$)

Si A es el suceso de nuestro interés, cuando ocurre A decimos que se ha producido un éxito y si no ocurre A decimos que se ha producido un fracaso.

$$P(A) = p > 0$$

$$P(\bar{A}) = 1 - P(A) = 1 - p = q > 0 \qquad p + q = 1$$

Ejemplos de experimentos de Bernouilli son todos aquellos en los que se establece una dicotomía entre dos sucesos como:

- Un artículo puede ser defectuoso o no serlo
- La estatura de un individuo puede llegar a 2m o no
- Un individuo puede comprar un producto o no comprarlo

Distribución de Bernouilli

La variable aleatoria X de Bernouilli cuenta el número de exitos en un experimento de Bernouilli

$$X \begin{cases} 0 : \text{fracaso, ocurre } \bar{A}, \ P(\bar{A}) = 1 - p = q \\ 1 : \text{éxito, ocurre } A, \ P(A) = p \end{cases}$$

La función de probabilidad de Bernouilli es:

x	$P(X=x)$
0	q
1	p

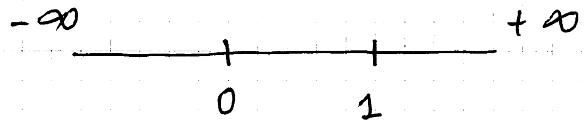
Esta tabla la podemos recoger en la expresión:

$$P(X=x) = p^x (1-p)^{1-x} = p^x \cdot q^{1-x}, \quad x = 0, 1.$$

Abreviadamente se representa por $X \rightsquigarrow b(p)$ a una

distribución de Bernouilli de parámetro p.
Veamos cual es la expresión de la función de distribución del experimento de Bernouilli. Esta función viene dada por:

$F(x) = P[X \leq x]$. Como X toma tan sólo los valores 0 y 1, el campo de variación de X viene dado por los siguientes intervalos:



Entonces:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ q & \text{si } 0 \leq x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

Veamos ahora las características de la distribución de Bernouilli.

1º Función generatriz de momento

$$\phi_X(t) = E[e^{tX}] = \sum_{x=0}^{1} e^{tx} \cdot P[X=x] = e^{t\cdot 0} \cdot P[X=0] + e^{t\cdot 1} \cdot P[X=1]$$

$\boxed{\phi_X(t) = q + p \cdot e^t}$, esta es la función generatriz de momento

2º Media

$$E[X] = \left.\frac{\partial \phi_X(t)}{\partial t}\right|_{t=0} = \left. p \cdot e^t \right|_{t=0} = p \Rightarrow \boxed{E[X] = p}$$

2º Varianza

$$\sigma^2 = E[X^2] - E[X]^2$$

$$E[X^2] = \left.\frac{\partial^2 \phi_X(t)}{\partial t^2}\right|_{t=0} = \left.\frac{\partial}{\partial t}\left(p \cdot e^t\right)\right|_{t=0} = \left. p \cdot e^t \right|_{t=0} = p$$

Por tanto

$$\sigma^2 = p - p^2 = p(1-p) = p \cdot q \Rightarrow \boxed{\sigma^2 = p \cdot q}$$

Veamos algunos ejemplos de distribución de Bernouilli.
Una persona dedicada a la venta de seguros de vida, realiza visitas a posibles clientes. Se sabe que en el 60% de las entrevistas tiene éxito y contrata un seguro. Defina la variable aleatoria asociada a este experimento y obtenga su media y su varianza.

Este es un experimento de Bernouilli ya que tan sólo se pueden dar dos sucesos: contratar un seguro (exito) = A, no contratar el seguro (fracaso) $\bar{A}$.
Ademas: $$P(A) = \frac{60}{100} = \frac{6}{10} = \frac{3}{5} \Rightarrow P(\bar{A}) = \frac{2}{5}$$

Si X = nº de seguros contratados en la entrevista.

x	$P[X=x]$
0	2/5
1	3/5

La función de probabilidad es $P[X=x] = \left(\frac{3}{5}\right)^x \cdot \left(\frac{2}{5}\right)^{1-x}$

La función generatriz de momentos es:
$$\phi_X(t) = E[e^{xt}] = e^{0t}\cdot\frac{2}{5} + e^{1\cdot t}\cdot\frac{3}{5} = \frac{2}{5} + \frac{3}{5}e^t$$

La media es:
$$\mu = E[X] = \left.\frac{\partial \phi_X(t)}{\partial t}\right|_{t=0} = \left.\frac{3}{5}e^t\right|_{t=0} = \frac{3}{5}$$

La varianza es $\sigma^2 = E[X^2] - E[X]^2$
$$E[X^2] = \left.\frac{\partial^2 \phi_X(t)}{\partial t^2}\right|_{t=0} = \left.\frac{\partial}{\partial t}\left(\frac{3}{5}e^t\right)\right|_{t=0} = \left.\frac{3}{5}e^t\right|_{t=0} = \frac{3}{5}$$

Luego $\sigma^2 = \frac{3}{5} - \left(\frac{3}{5}\right)^2 = \frac{3}{5} - \frac{9}{25} = \frac{15}{25} - \frac{9}{25} = \frac{6}{25}$

Vamos a pasar a otro ejemplo

Otro ejemplo de distribución de Bernouilli.
Entre los empleados de un taller, un tercio de ellos son especialistas. Elegido un empleado al azar se observa si es o no especialista. Defina la variable aleatoria asociada a este experimento aleatorio y obtenga su media y su varianza

De nuevo se trata de un experimento de Bernouilli. Se pueden dar tan sólo dos sucesos dicotómicos al elegir un empleado al azar:
A: el trabajador es especialista $P(A) = \frac{1}{3}$

$\bar{A}$: el trabajador no es especialista $P(\bar{A}) = \frac{2}{3}$
Si X = nº de especialistas al elegir un trabajador al azar

x	$P[X=x]$
0	$\frac{2}{3} = q$
1	$\frac{1}{3} = p$

La función de probabilidad es $P[X=x] = \left(\frac{1}{3}\right)^x \cdot \left(\frac{2}{3}\right)^{1-x}$ $x = 0,1$.
Según hemos visto en teoría:
- La media es $E[X] = \mu = p = \frac{1}{3}$

- La varianza es $\sigma^2 = p \cdot q = \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}$

Tercer ejemplo del experimento de Bernouilli.
Se lanzan al aire un par de dados y se observa si se obtiene o no una suma par de puntos. Defina la variable aleatoria asociada a este experimento aleatorio y obtenga su media y su varianza.

Los sucesos que se pueden dar en este experimento son dicotómicos: A = suma par (éxito) $P(A) = \frac{1}{2}$, $\bar{A}$ = suma impar (fracaso) $P(\bar{A}) = \frac{1}{2}$.
Si la variable es X = nº de lanzamientos en los que la suma es par

x	$P[X=x]$
0	$\frac{1}{2} = q$
1	$\frac{1}{2} = p$

La función de probabilidad es $P[X=x] = \left(\frac{1}{2}\right)^x \cdot \left(\frac{1}{2}\right)^{1-x} = \frac{1}{2}$

La media es $E[X] = \mu = p = \frac{1}{2}$

La varianza es $Var[X] = \sigma^2 = p \cdot q = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

Cuarto ejemplo del experimento de Bernouilli

En un lote con N artículos fabricados por una empresa hay N_1 defectuosos. Se toma un artículo al azar y se observa si es defectuoso o no. Defina la variable aleatoria asociada a este experimento y obtenga su media y su varianza

Elegido un artículo al azar o es defectuoso (éxito) A, o bien no lo es (fracaso) $\bar{A}$.

$$P(A) = \frac{N_1}{N} \qquad P(\bar{A}) = 1 - \frac{N_1}{N}$$

Sea la variable X = nº de artículos defectuosos al elegir uno al azar. Entonces X puede tomar valores 0 ó 1.

x	$P[X=x]$
0	$1 - \frac{N_1}{N} = q$
1	$\frac{N_1}{N} = p$

La función de probabilidad es:

$$P[X=x] = \left(\frac{N_1}{N}\right)^x \cdot \left(1 - \frac{N_1}{N}\right)^{1-x}$$

La media es $\mu = E[X] = p = \frac{N_1}{N}$

La varianza es:

$$\sigma^2 = Var[X] = p \cdot q = \left(\frac{N_1}{N}\right)\left(1 - \frac{N_1}{N}\right)$$

Distribución binomial

El experimento en la distribución binomial consiste en la repetición n-veces de uno de Bernouilli, siendo cada uno de estas repeticiones independiente de las demás.

La variable aleatoria X de la distribución binomial es el nº de éxitos obtenidos al repetir n-veces un experimento de Bernouilli:

Si $X_i \rightsquigarrow B(p)$, la variable X de la binomial es

$$X = \sum_{i=1}^{n} X_i$$

Vamos a obtener la función de probabilidad $P[X=x]$. Hemos de tener en cuenta las repeticiones en donde hay x éxitos y n-x fracasos. Una de estas repeticiones es: $R_1 = A \cap A \cap \overset{x}{\cdots} \cap A \cap \bar{A} \cap \bar{A} \cap \overset{n-x}{\cdots} \cap \bar{A}$ (con esto queremos indicar que hemos obtenido exito en la 1ª y en la 2ª repetición y en la xª, obteniendo fracaso en la siguiente, y en la siguiente y en la última). Como las repeticiones son independientes unas de otras:

$P(R_1) = P(A \cap A \cap \overset{x}{\cdots} \cap A \cap \bar{A} \cap \bar{A} \cap \overset{n-x}{\cdots} \cap \bar{A})$. Luego:

$P(R_1) = P(A) \cdot P(A) \overset{x}{\cdots} P(A) \cdot P(\bar{A}) \cdot P(\bar{A}) \overset{n-x}{\cdots} P(\bar{A})$, es decir

$$\boxed{P(R_1) = p^x \cdot q^{n-x}}$$

Nos preguntamos ahora cuantas repeticiones hay en las que obtengamos x exitos y n-x fracasos. Otra repetición es: $R_2 = (A \cap A \cap \overset{x}{\cdots} \cap \bar{A} \cap A \cap \bar{A} \cap \overset{n-x}{\cdots} \cap \bar{A})$. ¿Qué ha cambiado entre R_1 y R_2? El orden, tan sólo el orden, ya que tanto en R_1 como en R_2 hay un número de $A = x$ y un número de $\bar{A} = n-x$. Como en cada repetición los elementos A y $\bar{A}$ se pueden repetir el nº de repeticiones se obtiene con permutaciones con repetición

$$PR_n^{x,n-x} = \frac{n!}{x!\,(n-x)!} = \binom{n}{x}$$

El que $PR_n^{x, n-x} = \binom{n}{x}$ es pura casualidad, conceptualmente son dos cosas distintas. Hemos probado que el nº de repeticiones en donde hay x éxitos y n-x fracasos es:

$$k = \binom{n}{x}$$

Entonces las repeticiones en donde hay x éxitos y n-x fracasos son: $R_1, R_2, \dots R_K$. Por tanto:

$P[X=x] = P[R_1 \cup R_2 \cup \dots \cup R_K]$. Cada una de estas repeticiones es incompatible con las otras, es decir $R_i \cap R_j = \phi$ si $i \neq j$ y además la probabilidad de cada una de estas repeticiones es: $P(R_i) = p^x q^{1-x}$ $i = 1 \dots K$. Luego:

$$P[X=x] = P(R_1) + P(R_2) + \dots + P(R_K) = K \cdot P[R_1] = \binom{n}{x} p^x \cdot q^{(1-x)}$$

La función de probabilidad de la binomial es

$$\boxed{P[X=x] = \binom{n}{x} p^x \cdot q^{(1-x)}} \qquad x = 0, 1, 2, \dots$$

La función de distribución es $F(x) = P[X \leq x]$

$F(x) = \sum_{x_i \leq x} P[X=x_i] = \sum_{x_i \leq x} \binom{n}{x_i} \cdot p^{x_i} \cdot q^{1-x_i}$ esta función no es fácil de manejar por eso existen tablas para calcularla.

Si la variable X corresponde a una distribución binomial de parámetros n y p, abreviadamente se escribe $X \rightsquigarrow B(n, p)$

Características de la distribución binomial

1º Relación entre la distribución de Bernouilli y binomial

$B(1, p) = b(p)$: la binomial con una repetición es la distribución de Bernouilli

2º Función generatriz de momentos.

Recordemos que si $x_1, x_2, \dots x_n$ son independientes y $X = \sum_{i=1}^{n} X_i$, entonces la función generatriz de momentos de X vale:

$$\phi_X(t) = \phi_{X_1}(t) \cdot \phi_{X_2}(t) \cdot \dots \cdot \phi_{X_n}(t)$$

Al ser X_i la variable de una distribución de Bernouilli, hemos visto que $\phi_{X_1}(t) = \phi_{X_2}(t) = \dots = \phi_{X_n}(t) = q + p \cdot e^t$

Por tanto la función generatriz de momentos es

$$\phi_X(t) = (q + pe^t)^n$$

3º Media y varianza

La media es $E[X] = \left.\frac{d\phi_X(t)}{dt}\right|_{t=0} = n \cdot (q+pe^t)^{n-1} \cdot pe^t \Big|_{t=0}$

$E[X] = n \cdot p(q+p)^{n-1}$. Como $q+p=1 \Rightarrow E[X] = n \cdot p$

La varianza es $Var[X] = \sigma^2 = E[X^2] - E[X]^2$

$$E[X^2] = \left.\frac{\partial^2 \phi_X(t)}{\partial t^2}\right|_{t=0} = n \cdot (n-1)(q+pe^t)^{n-2} \cdot p^2 e^{2t} + n(q+pe^t)^{n-1} \cdot pe^t \Big|_{t=0}$$

$E[X^2] = n(n-1)p^2 + np = n^2p^2 - np^2 + np = n^2p^2 + np(-p+1)$;

$E[X^2] = n^2p^2 + npq$

Entonces $\sigma^2 = n^2p^2 + npq - (np)^2 = npq$

4º Propiedad reproductiva

Sean $X_1, X_2, \dots X_n$ v.a. independientes tales que $X_i \rightsquigarrow B(n_i, p)$. Entonces:

$$X = \sum_{i=1}^{n} X_i \rightsquigarrow B\left(\sum_{i=1}^{n} n_i, p\right)$$

5ª En el mismo experimento aleatorio se define:

X : número de éxitos $X \rightsquigarrow B(n, p)$

Y : número de fracasos $Y \rightsquigarrow B(n, q)$. Entonces:

$P(X = x) = P[Y = n - x]$

Ejemplos de la distribución binomial

1º Entre los empleados de un taller un tercio son especialistas. Si elegimos 4 de ellos con reemplazamiento:
a) ¿Cuál es la probabilidad de que dos de ellos sean especialistas?
b) ¿Cuál es la probabilidad de que como máximo dos de ellos sean especialistas?

Sea X = número de especialistas cuando se eligen 4 empleados al azar. Nos piden $P(X=2)$. Cada una de las selecciones es una repetición de un experimento de Bernouilli con probabilidad de exito $p = \frac{1}{3}$ (probabilidad de escoger a un especialista). Además son repeticiones independientes ya que la selección se realiza con reemplazamiento. (el resultado de cada una de ellas no depende de las anteriores). Por tanto $X \rightsquigarrow B(4, \frac{1}{3})$

La probabilidad $P[X=2]$ se puede calcular de tres formas:

a) Utilizando la distribución de probabilidad binomial

$$P[X=x] = \binom{n}{x} p^{x} \cdot q^{1-x}$$

$$P[X=2] = \binom{4}{2} \left(\frac{1}{3}\right)^2 \cdot \left(\frac{2}{3}\right)^2 = 0'2963$$

b) Utilizando las tablas estadisticas de la función de probabilidad binomial:
Como $n = 4$, $x = 2$ y $p = \frac{1}{3} \Rightarrow P[X=2] = 0'2963$

c) Utilizando las tablas estadisticas de la función de distribución:

$$P[X=2] = P[X \leq 2] - P[X \leq 1] = F(2) - F(1) = 0'8889 - 0'5926$$

$$P[X=2] = 0'2963$$

Ejemplo: Se envían 20 invitaciones a los representantes estudiantiles para asistir a una conferencia. De experiencias anteriores se sabe que la probabilidad de aceptar una invitación es de 0'8. Si las decisiones de aceptar estas invitaciones son independientes, determinar la probabilidad de que como máximo 17 estudiantes acepten la invitación.

Tomamos la variable estadística:

$$X_i = \begin{cases} 1 & \text{si el estudiante } i \text{ acepta la invitación} \\ 0 & \text{si el estudiante } i \text{ no acepta la invitación} \end{cases}$$

Entonces $X_i \rightsquigarrow B(p)$ $\quad i = 1 \cdots 20 \quad$ $p = 0'8$

Sea $X = \sum_{i=1}^{20} X_i$. Entonces $X \rightsquigarrow B(n,p) = B(20, 0'8)$

Nos piden $P(X \leq 17) = 1 - P[X > 17] = 1 - P(X=18) - P(X=19) -$
$- P(X=20)$.

Al ser $X \rightsquigarrow B(20, 0'8) \Rightarrow P[X = x] = \binom{20}{x} 0'8^{x} \cdot 0'2^{20-x}$

Luego:

$$P[X \leq 17] = 1 - \binom{20}{18} 0'8^{18} \cdot 0'2^{2} - \binom{20}{19} \cdot 0'8^{19} \cdot 0'2 - \binom{20}{20} 0'8^{20} = 0'794$$

Otra forma de resolver este ejemplo sería:

Y = número de representantes que no aceptan la invitación entre los 20 invitados. Entonces:

$Y \rightsquigarrow B(20, 0'2) \qquad X + Y = 20 \Rightarrow X = 20 - Y$

$$P[X \leq 17] = P[20 - Y \leq 17] = P[Y \geq 3] = 1 - P[Y < 3] =$$

$$= 1 - P[Y \leq 2] = 1 - \underbrace{F(2)}_{\text{tablas}} = 1 - 0'2061 = 0'7939.$$

Ejemplo: Se lanzan dos dados diez veces al aire, ¿cuál es la probabilidad que en más de la mitad de las ocasiones se obtenga una suma par de puntos?

Tomamos la variable aleatoria

$$X_i = \begin{cases} 0 \text{ si en el lanzamiento } i \text{ no se obtiene suma par} \\ 1 \text{ si en el lanzamiento } i \text{ se obtiene suma par} \end{cases}$$

Entonces $X_i \rightsquigarrow B(p=\frac{1}{2}) \quad i=1 \cdots 10$

Consideramos la v.a. $X = \sum_{i=1}^{20} X_i$. Nos preguntan

$$P[X>5] = 1-P[X \leq 5] = 1-\underset{\text{tablas}}{F(5)} = 1-0'6230 = 0'377.$$

DISTRIBUCION GEOMETRICA O DE PASCAL

Experimento aleatorio: sucesión de repeticiones independientes de un experimento de Bernouilli con probabilidad de éxito p. Las repeticiones cesan cuando aparece el primer exito.

Variable aleatoria: es el nº de fracasos que tienen lugar hasta que aparece el primer éxito.
La función de probabilidad $P[X=x]$ indica que hay x fracasos antes del primer éxito, la sucesión de repeticiones de experimentos de Bernouilli sería:

$$\underbrace{\bar{A} \cap \bar{A} \cap \cdots \cap \bar{A}}^{x} \cap A \Rightarrow P[X=x] = P[\underbrace{\bar{A} \cap \bar{A} \cap \cdots \cap \bar{A}}^{x} \cap A] \Rightarrow$$

$$\Rightarrow P[X=x] = \underbrace{P(\bar{A}) \cdot P(\bar{A}) \cdots P(\bar{A})}^{x} \cdot P(A) \Rightarrow P[X=x] = q^x \cdot p$$

La función de probabilidad es $P[X=x] = p \cdot q^x \quad x = 1, 2, \ldots$

Cuando X se distribuye según una distribución geométrica de parámetro p lo indicamos como:

$$X \rightsquigarrow G(p)$$

Una v.a. relacionada con la geométrica es $Y = X+1$, es decir el número de repeticiones de un experimento de Bernouilli hasta que aparece el primer éxito. Entonces:

$$P[Y=y] = P[X+1=y] = P[X=y-1] = p\cdot q^{y-1} \quad y=1,2,\dots$$

Características de la distribución geométrica o de Pascal:

1º Función de distribución

$$F(x) = P[X \leq x] = \sum_{x_i \leq x} P[X=x_i] = \sum_{x_i=0}^{x} p\cdot q^{x_i} = p\cdot\sum_{x_i=0}^{x} q^{x_i}$$ (se trata de la suma de x-términos de una progresión geométrica de razón q, se halla mediante la expresión $\frac{a_1 - a_n\cdot r}{1-r}$) es decir $F(x) = p\cdot\frac{1-q^x\cdot q}{1-q} = \cancel{p}\cdot\frac{1-q^{x+1}}{\cancel{p}}$

luego $F(x) = 1 - q^{x+1}$

2º Función generatriz de momentos

La función generatriz de momentos es

$$\phi_x(t) = E[e^{tx}] = \sum_{x=0}^{\infty} e^{tx}\cdot P[X=x] = \sum_{x=0}^{\infty} e^{tx}\cdot q^x\cdot p$$

$\phi_x(t) = p\cdot\sum_{x=0}^{\infty}(q\cdot e^t)^x$. Se trata pues de sumar los infinitos términos de una progresión geométrica de razón $r = q\cdot e^t$. Para que esta suma se pueda efectuar ha de ser $|r| = q\cdot e^t < 1 \Rightarrow e^t < \frac{1}{q} \Rightarrow t < -\ln q$

En este caso la suma vale $\frac{a_1}{1-r}$, es decir:

$$\phi_x(t) = p\cdot\frac{1}{1-q\cdot e^t} \quad \text{si } t < -\ln q.$$

luego $\phi_x(t) = \frac{p}{1-q\cdot e^t}$ si $t < -\ln q$

3º Media

$$E[x] = \left.\frac{d\,\phi_x(t)}{dt}\right|_{t=0} = \left.\frac{p\cdot q\cdot e^t}{(1-e^t q)^2}\right|_{t=0} = \frac{p\cdot q}{p^2} = \frac{q}{p}$$

Luego $E[x] = \dfrac{q}{p}$

4ª Varianza

$$Var[x] = \sigma^2 = E[x^2] - E[x]^2$$

$$E[x^2] = \left.\frac{d^2\phi_x(t)}{dt^2}\right|_{t=0} = \left.\frac{p\cdot q\cdot e^t(1-q\cdot e^t)^2 - pq e^t\cdot 2(1-e^t q)(-e^t q)}{(1-e^t q)^4}\right|_{t=0}$$

$$E[x^2] = \left.\frac{p\cdot q\cdot e^t(1-q\cdot e^t) + 2pq^2\cdot e^{2t}}{(1-q\cdot e^t)^3}\right|_{t=0} = \frac{p^2q+2q^2p}{p^3} = \frac{pq+2q^2}{p^2}$$

$$E[x^2] = \frac{q(p+2q)}{p^2} = \frac{q(1-q+2q)}{p^2} = \frac{q(1+q)}{p^2}$$

Luego: $\sigma^2 = \dfrac{q(1+q)}{p^2} - \left(\dfrac{q}{p}\right)^2 = \dfrac{q+q^2-q^2}{p^2} = \dfrac{q}{p^2}$

$$\sigma^2 = \frac{q}{p^2}$$

5ª No se verifica la propiedad reproductiva

Sean $X_1, X_2, \dots X_n \rightsquigarrow G(p)$. Sea $X = \sum_{i=1}^{n} X_i$. Entonces no tiene porqué verificarse $X \rightsquigarrow G(p)$, en general $X \not\rightsquigarrow G(p)$

Veamos un ejemplo.

Lanzamos un dado hasta que aparezca el número 6. ¿Cuál es la probabilidad de que el primer 6 aparezca en el cuarto lanzamiento?

Sea X_i la siguiente variable de Bernouilli

$$X_i = \begin{cases} 0 \text{ si no sale } 6 & q = \frac{5}{6} \\ 1 \text{ si sale } 6 & p = \frac{1}{6} \end{cases}$$

Entonces $X_i \rightsquigarrow G(\frac{1}{6})$

Consideremos X = nº de fracasos hasta que repitiendo el experimento de Bernouilli aparece un 6.

Entonces $X \rightsquigarrow G(\frac{1}{6})$ y nos piden $P[X=3]$

Como la función de probabilidad de la distribución geométrica $G(p)$ viene dada por

$$P[X=x] = p \cdot q^x = \left(\frac{1}{6}\right) \cdot \left(\frac{5}{6}\right)^x$$

$$P[X=3] = \left(\frac{1}{6}\right) \cdot \left(\frac{5}{6}\right)^3 = \frac{5^3}{6^4} = 0'0965$$

DISTRIBUCION BINOMIAL NEGATIVA

Experimento aleatorio: se repiten experimentos de Bernouilli independientes hasta que aparece el r-ésimo exito.

Variable aleatoria: número de fracasos hasta que aparece el r-ésimo exito.

Vamos a obtener la función de probabilidad $P[X=x]$

Han aparecido x fracasos antes del r-ésimo éxito.

Formemos uno de estos sucesos:

$$G_1 = \overbrace{\bar{A} \cap \bar{A} \cap \ldots \cap \bar{A}}^{x} \cap \overbrace{A \cap A \cap \ldots \cap A \cap A}^{r}$$

↓el último suceso siempre es un éxito

$$P(G_1) = \overbrace{P(\bar{A}) \cdot P(\bar{A}) \cdot \ldots \cdot P(\bar{A})}^{x} \cdot \overbrace{P(A) \cdot P(A) \cdot \ldots \cdot P(A)}^{r} = q^x \cdot p^r$$

siendo $q = P(\bar{A})$ (probabilidad de fracaso) $= 1-p$

$p = P(A)$ (probabilidad de éxito)

Otro suceso de los que nos interesan es:

$$G_2 = \bar{A} \cap \bar{A} \cap \overset{x}{\ldots} \cap A \cap \bar{A} \cap A \cap \ldots \cap A \cap A$$ → el último suceso siempre es un éxito

$$P(G_2) = q^x \cdot p^r$$

¿Que diferencia a G_1 y G_2? Tan sólo el orden, en cualquiera de estos sucesos aparecen x repeticiones de $\bar{A}$

y r repeticiones de A.

$[X=x]$ es el nº de fracasos obtenidos antes de que aparezcan r-exitos. ¿Cuántos sucesos hay de este tipo? Como la última repetición es un éxito:

$$PR_{x+r-1}^{x\ ,r-1} = \frac{(x+r-1)!}{x!\ (r-1)!} = \binom{x+r-1}{x} = K$$

De nuevo es casualidad que $PR_{x+r-1}^{x,\ r-1} = \binom{x+r-1}{x}$

luego $[X=x] = G_1 \cup G_2 \cup \cdots \cup G_K$ y como estos G_i son incompatibles dos a dos:

$P[X=x] = P[G_1 \cup G_2 \cup \cdots \cup G_K] = P(G_1)+P(G_2)+\cdots+P(G_K)$.

Como $P(G_1) = P(G_2) = \cdots = P(G_K) = p^r \cdot q^x$

$P[X=x] = K \cdot p^r \cdot q^x$. Como $K = \binom{x+r-1}{x}$ la función de probabilidad es:

$$P[X=x] = \binom{x+r-1}{x} p^r \cdot q^x$$

En esta distribución binomial negativa una variable relacionada con $X=$ nº fracasos hasta obtener r exitos es $Y=$ nº ensayos hasta obtener r-éxitos. Entonces $Y=X+r$ y su función de probabilidad es:

$$P[Y=y] = P[X+r=y] = P[X=y-r] = \binom{y-r+r-1}{y-r} p^r \cdot q^{y-r}$$

$$\Rightarrow P[Y=y] = \binom{y-1}{y-r} p^r \cdot q^{y-1}$$

Características de la distribución binomial negativa

1º Relación entre la binomial negativa y la geométrica

$BN(1,p) = G(p)$

Si X se distribuye según una binomial negativa

$X \rightsquigarrow BN(r,p)$

2º Función generatriz de momentos

La distribución binomial negativa es la repetición r veces de una variable geométrica de parámetro p. Si $X_i \rightsquigarrow G(p)$ es el nº de fracasos antes del primer éxito $X = \sum_{i=1}^{r} X_i$ representa el número de fracasos hasta obtener r éxitos. Por tanto $X \rightsquigarrow BN(r,p)$. Como las X_i son independientes y su función generatriz de momentos es $\phi_{X_i}(t) = \frac{p}{1-qe^t}$ se tiene que:

$$\phi_X(t) = \phi_{X_1}(t) \cdot \phi_{X_2}(t) \cdots \overset{r}{\smile} \cdots \phi_{X_r}(t) = \left(\frac{p}{1-qe^t}\right)^r$$

La función generatriz de momentos de la binomial negativa es:

$$\phi_X(t) = \left(\frac{p}{1-qe^t}\right)^r$$

3º Media

$$E[X] = \left.\frac{d\phi_X(t)}{dt}\right|_{t=0} = r \cdot \left(\frac{p}{1-qe^t}\right)^{r-1} \cdot \frac{pq \cdot e^t}{(1-qe^t)^2} = \left.\frac{r \cdot p^r}{(1-q e^t)^{r+1}} \cdot qe^t\right|_{t=0}$$

$$E[X] = r\frac{p^r}{(1-q)^{r+1}} \cdot q = r\frac{p^r \cdot q}{p^{r+1}} = r\frac{q}{p} \Rightarrow E[X] = r\frac{q}{p}$$

4º Varianza

$$\sigma^2 = Var[X] = E[X^2] - E[X]^2$$

$$E[X^2] = \left.\frac{d^2\phi_X(t)}{dt^2}\right|_{t=0} = \frac{d}{dt}\left[\frac{r \cdot p^r}{(1-q\cdot e^t)^{r+1}} \cdot q \cdot e^t\right]_{t=0}$$

$$E[X^2] = \left.\frac{r \cdot p^r \cdot q \cdot e^t (1-q\cdot e^t)^{r+1} + r \cdot p^r \cdot q^2 \cdot e^{2t} \cdot (r+1) \cdot (1-q\cdot e^t)^r}{(1-qe^t)^{2r+2}}\right|_{t=0}$$

$$E[x^2] = \left. \frac{r p^r \cdot q \cdot e^t (1 - q \cdot e^t) + r(r+1) \cdot p^r \cdot q^2 \cdot e^{2t}}{(1 - q \cdot e^t)^{r+2}} \right|_{t=0}$$

$$E[x^2] = \frac{r p^{r+1} \cdot q \cdot + r(r+1) \cdot p^r \cdot q^2}{p^{r+2}} = \frac{r p \cdot q + r(r+1) \cdot q^2}{p^2}$$

$$E[x^2] = \frac{r p q + r^2 q^2 + r q^2}{p^2} \quad \text{. Luego :}$$

$$\sigma^2 = \frac{r p q + r^2 q^2 + r q^2}{p^2} - \left(r \frac{q}{p}\right)^2 = \frac{r p q + r \cdot q^2 + \cancel{r^2 q^2} - \cancel{r^2 q^2}}{p^2}$$

$$\sigma^2 = \frac{r p q + r \ q^2}{p^2} = \frac{r(1-q) \cdot q + r \ q^2}{p^2} = \frac{r q - \cancel{r q^2} + \cancel{r q^2}}{p^2}$$

Por tanto $\sigma^2 = \dfrac{r q}{p^2}$

5º La binomial negativa verifica la propiedad reproductiva

Sean $X_i \rightsquigarrow BN(r_i, p)$ $i = 1 \dots n$, X_i independientes.
Sea $X = \sum_{i=1}^{n} X_i$. Entonces:

$$X \rightsquigarrow BN\left(\sum_{i=1}^{n} r_i, p\right)$$

Ejemplo: En un departamento de control de calidad se inspeccionan las unidades terminadas que provienen de una linea de ensamblaje. Se piensa que la proporción de unidades defectuosas es de 0'05.

a) ¿cuál es la probabilidad de que la vigésima unidad inspeccionada sea la segunda que se encuentra defectuosa?

Si X = nº de unidades no defectuosas hasta que se encuentran dos defectuosas, considerando que ser defectuosa es el exito ($p = 0'05$) se trata de una binomial negativa $X \rightsquigarrow BN(2, 0'05)$. Nos piden la probabilidad que haya 18 no defectuosas hasta que se obtienen 2 defectuosas

$$P[X=x] = \binom{x+r-1}{x} p^r \cdot q^x \qquad x=18 \quad r=2 \quad p=0'05 \quad q=0'95$$

Por lo tanto:

$$P[X=18] = \binom{18+2-1}{18} \cdot (0'05)^2 \cdot (0'95)^{18} =$$

$$P[X=18] = \binom{19}{18} \cdot (0'05)^2 \cdot (0'95)^{18} = 0'0189$$

b) ¿Cuantas unidades se tienen que inspeccionar por termino medio, hasta encontrar cuatro defectuosas?

Si X = nº de unidades no defectuosas hasta encontrar 4 defectuosas $\Rightarrow X \rightsquigarrow BN(4, 0'05)$

Si Y = nº de unidades inspeccionadas hasta encontrar 4 defectuosas $\Rightarrow Y = X+4$

$$E[Y] = E[X+4] = 4 + E[X]$$

Como $E[X] = \dfrac{r \cdot q}{p} = \dfrac{4 \cdot 0'95}{0'05} = 4 \cdot 19 = 76$

Por tanto $E[Y] = 4 + 76 = 80$ unidades

c) Calcule la desviación típica del número de unidades que se deben inspeccionar hasta encontrar la cuarta defectuosa.

$$Y = X+4 \Rightarrow \sigma(Y) = \sigma(X+4) = \sigma(X) = \sqrt{\sigma^2(X)}$$

$$\sigma(Y) = \sqrt{\frac{r \cdot q}{p^2}} = \sqrt{\frac{4 \cdot 0'95}{0'05^2}} = 38'987.$$

Distribución Hipergeométrica.

Experimento: Sea una población con N individuos de los cuales N_1 tienen una determinada característica y $N_2 = N - N_1$ no la tienen.

Si consideramos éxito el tener la característica entonces $p = \frac{N_1}{N}$ y $q = 1 - p = \frac{N_2}{N}$

extraemos una muestra de tamaño n de la anterior población sin reemplazamiento (esto quiere decir que examinamos los individuos de la muestra no uno por uno devolviéndolos a la población tras su examen, sino en conjunto es como examinamos el grupo. Si hubiera reemplazamiento se trataría de una distribución binomial).

Variable aleatoria: número de éxitos obtenidos al examinar la muestra de tamaño n. Veamos cual es la función de probabilidad $P[X=x]$. Si en la muestra de tamaño n hay x éxitos y $n-x$ fracasos, la forma en que podemos obtener x exitos viene dado por $\binom{N_1}{x}$ y los $n-x$ fracasos viene dado por $\binom{N_2}{n-x}$

Cada exito se puede combinar con un fracaso para formar un suceso favorables. Luego:

$$\text{Casos favorables} = \binom{N_1}{x} \cdot \binom{N_2}{n-x} = \binom{Np}{x} \cdot \binom{Nq}{n-x}$$

Los casos posibles serían todas las muestras posibles que se pueden formar con los N individuos de la población, las muestras de tamaño n claro está, los casos posibles son $\binom{N}{n}$

Luego

$$P[X=x] = \frac{\binom{Np}{x} \cdot \binom{Nq}{n-x}}{\binom{N}{n}}$$ esta es la función de probabilidad de la distribución hipergeométrica.

Para conocer la distribución hipergeométrica necesitamos los parámetros N, n y p. Si X se distribuye según una distribución hipergeométrica se escribe $X \rightsquigarrow H(N, n, p)$

Veamos como varía la v. a X. Se tiene que

$$\left.\begin{array}{l} X \leq n \\ X \leq N_1 \end{array}\right\} \Rightarrow X \leq \min (n, N_1)$$

$$\left.\begin{array}{l} X \geq 0 \\ n - X \leq N_2 \end{array}\right\} \Rightarrow \begin{array}{l} X \geq 0 \\ n - N_2 \leq X \end{array} \Rightarrow X \geq \max \{0, n - N_2\}$$

Luego: $\max \{0, n - N_2\} \leq X \leq \min \{n, N_1\}$

Características de la distribución hipergeométrica.

- La función generatriz de momentos en una hipergeométrica es muy difícil de obtener, lo mismo ocurre con la deducción de la esperanza y la varianza.

a) Esperanza de una distribución hipergeométrica

$$E[X] = n \cdot p$$

b) Varianza de una distribución hipergeométrica

$$Var[X] = \sigma^2 = n \cdot p \cdot q \cdot \frac{N-n}{N-1}$$

c) Relación entre la distribución hipergeométrica y la binomial.

Sea $X \rightsquigarrow H(N, n, p)$ tal que N es muy grande en comparación con n, o de otra forma $\frac{n}{N} \to 0$

$\lim_{N \to \infty} \frac{n}{N} = 0$. Entonces:

$$\lim_{N \to \infty} H(N, n, p) = B(n, p)$$

Cuando la población es muy grande $(N \to \infty)$ la hipergeométrica se aproxima a la binomial

Veamos ejemplos de distribución hipergeométrica.

1º En un pedido de 100 pares de zapatos se sabe que 10 pares son defectuosos. Si escogemos a la vez 8 pares, ¿ cual es la probabilidad de que todos esten en perfectas condiciones ?

Se trata de una distribución hipergeométrica ya que se escoge la muestra sin devolución

X : nº de éxitos (que no haya defecto) en una muestra de 8 = n.

$X \rightsquigarrow H(100, 8, 0'9)$

Nos estan preguntando $P[X=8] = \dfrac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}}$;

$$P[X=8] = \frac{\binom{100\cdot 0'9}{8}\cdot\binom{100\cdot 0'1}{0}}{\binom{100}{8}} = \frac{\binom{90}{8}\cdot\binom{10}{0}}{\binom{100}{8}} = 0'4166$$

2º Una caja contiene 6 bolas blancas y 4 rojas. Se realiza una prueba consistente en extraer una bola y observar su color. Calcule la probabilidad de que, después de 5 pruebas se hayan escogido 3 bolas blancas si:
a) las bolas no se reemplazan. b) Las bolas se reemplazan.
Repita el ejercicio suponiendo que la caja contiene 6000 bolas blancas y 4000 bolas rojas.

a) Sea X = nº exitos (bolas blancas) que se han obtenido en la muestra de n = 5. En este caso $p = \frac{6}{10} = 0'6$. Como no hay reemplazamiento se trata de una distribución hipergeométrica, es decir $X \rightsquigarrow H(N, n, p) \Rightarrow$

$\Rightarrow X \rightsquigarrow H(10, 5, 0'6)$.

$P[X=x] = \dfrac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}}$; Como nos piden $P[X=3]$

$$P[X=3] = \frac{\binom{10\cdot 0'6}{3}\binom{10\cdot 0'4}{2}}{\binom{10}{5}} = \frac{\binom{6}{3}\cdot\binom{4}{2}}{\binom{10}{5}} = 0'476$$

b) Si hay reemplazamiento se distribuye según una binomial $B(n,p) = B(5, 0'6)$. Por tanto

$$P[X=3] = \binom{5}{3}\cdot 0'6^3 \cdot 0'2^2 = 0'3456$$

c) Supongamos que no hay reemplazamiento. Si el número de individuos de la población es $N = 6.000$ y el tamaño de la muestra es $n = 5$ se tiene que

$$\frac{5}{6000} \to 0 \quad \text{y por tanto} \quad H(N,n,p) \longrightarrow B(n,p)$$

Luego $P[X=3] = \binom{5}{3}\cdot 0'6^3\cdot 0'4^2 = 0'3456.$

Distribución de Poisson.

La variable aleatoria de Poisson recoge el número de acontecimientos independientes que ocurren a un ritmo constante sobre el tiempo o espacio.

Ejemplos de distribución de Poisson son:

- El número de clientes que llegan al banco por minuto
- El número de enfermos que llegan por hora al servicio de urgencias de un hospital
- El número de errores mecanográficos por página.

Si la v.a. X se distribuye según una distribución de Poisson de parámetro λ, su función de probabilidad es:

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

Abreviadamente se escribe $X \rightsquigarrow P(\lambda)$

Veamos ejemplos de la distribución de Poisson

Ejemplo: Supongamos que la demanda mensual de televisores de cierta marca sigue una distribución de Poisson de parámetro $\lambda = 10$ y un beneficio neto de 120 €.

a) ¿Cual es la probabilidad de que el beneficio neto mensual que obtenga un comerciante sea, al menos de 1440 €?

Si $X =$ nº de unidades demandadas

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

$$P[X = x] = \frac{10^x}{x!} e^{-10} \qquad x = 0,1,2,\ldots$$

Sea la variable $Y =$ beneficio neto. Entonces $Y = 10X$. Nos preguntan $P[Y \geq 1440] = P[120X \geq 1440] = P[X \geq 12] =$

$= 1 - P[X < 12] = 1 - P[X \leq 11] = 1 - F(11)$ donde $F(x)$ es la función de distribución de Poisson:

$$F(x) = \sum_{x_i \leq x} P[X = x_i]$$

$$F(11) = \sum_{x_i \leq 11} P[X = x_i] \underset{\text{Tablas}}{=} 0'6968$$

Luego $P[Y \geq 1440] = 1 - 0'6968 = 0'3032$

b) ¿Qué stock debe tener almacenado el comerciante a principio de mes para tener una probabilidad de 0'95 de satisfacer toda la demanda durante dicho mes?

Tenemos que calcular a de modo que

$P[X \leq a] \geq 0'95$. Si buscamos en las tablas ha de ser $a = 15$ ya que $F[X \leq 15] = 0'9513$

El stock ha de ser de 15 televisores.

Características de la distribución de Poisson:

a) Función generatriz de momentos.
La función generatriz de momentos es

$$\phi_x(t) = E[e^{tx}] = \sum_{x=0}^{\infty} e^{tx} \cdot \frac{\lambda^x}{x!} e^{-\lambda} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(e^t \cdot \lambda)^x}{x!}$$

$\phi_x(t) = e^{-\lambda} \cdot e^{\lambda \cdot e^t} \Rightarrow \boxed{\phi_x(t) = e^{\lambda(e^t - 1)}}$ esta es la función generatriz de momentos en una distribución de Poisson.

b) Media

$$E[x] = \left.\frac{d\,\phi_x(t)}{dt}\right|_{t=0} = \left.\lambda \cdot e^t \cdot e^{\lambda(e^t-1)}\right|_{t=0} = \lambda; \quad \boxed{E[x] = \lambda}$$

c) Varianza

$$Var[x] = \sigma^2 = E[x^2] - E[x]^2$$

$$E[x^2] = \left.\frac{d^2 \phi_x(t)}{dt^2}\right|_{t=0} = \left.\lambda \cdot e^t \cdot e^{\lambda(e^t-1)} + \lambda^2 e^{2t} \cdot e^{\lambda(e^t-1)}\right|_{t=0}$$

$$E[x^2] = \lambda + \lambda^2.$$

Luego: $\sigma^2 = (\lambda + \lambda^2) - \lambda^2 = \lambda \Rightarrow \boxed{\sigma^2 = \lambda}$

d) Propiedad reproductiva.
Sean $X_1, X_2, \dots X_n$ variables de Poisson de parámetros $\lambda_1, \lambda_2, \dots \lambda_n$ respectivamente $X_i \rightsquigarrow P(\lambda_i)$ $i = 1 \dots n$.
Si $X = \sum_{i=1}^{n} X_i$ entonces $X \rightsquigarrow P\left(\sum_{i=1}^{n} \lambda_i\right)$

e) Relación entre la distribución binomial y la de Poisson

Si n es muy grande y p es muy pequeño

$$B(n,p) \longrightarrow P(n \cdot p) \qquad \lambda = \text{media} = n \cdot p.$$

Ejemplo: Suponga que el número de llamadas que recibe una operadora entre las 9 y las 9'05 horas de cualquier día sigue una distribución de Poisson de parámetro 5. Determine la probabilidad de que después de dos días el nº total de llamadas recibidas en dicho intervalo sea 8.

Sea X_i = nº llamadas que recibe una operadora entre las 9 y las 9'05 en un día $i = 1, 2, 3, \ldots$

Entonces $X_i \rightsquigarrow P(5)$.

$X_1 \rightsquigarrow P(5)$

$X_2 \rightsquigarrow P(5)$

$X = X_1 + X_2$ llamadas recibidas en dos días. Según la propiedad reproductiva $X \rightsquigarrow P(5+5=10)$.

Nos piden $P[X=8] = \frac{10^8}{8!} e^{-10} \underset{\text{Tablas}}{=} F(8) - F(7) = 0'3328 - 0'2202$

$P[X=8] = 0'1126$.

Ejemplo: El nº medio de coches que llega a una estación de servicio es de 210/hora. Si dicha estación puede atender como máximo 10 coches/minuto

a) ¿Cuál es la probabilidad de que, en un minuto determinado, se atiendan 3 coches?

b) ¿Cuál es la probabilidad de que, en un minuto dado, haya dos coches guardando cola?

c) Determine la probabilidad de que, en un minuto dado, lleguen a la estación más coches de los que puede atender.

a) X: nº de coches que llegan a la estación en un minuto. Entonces

$E[X] = \frac{21\cancel{0}}{6\cancel{0}} = \frac{7}{2} = 3'5$. Entonces $X \rightsquigarrow P(\lambda)$ y como $\lambda = E[X] = 3'5$:

$X \rightsquigarrow P(3'5)$

Nos piden $P[X=3]$ $\qquad P[X=x] = \frac{\lambda^x}{x!} e^{-\lambda}$

$P[X=3] = \frac{(3'5)^3}{3!} e^{-3'5} = F(3)-F(2) = 0'5366 - 0'3208 = 0'2158$

$\downarrow$ Tablas

Luego $P[X=3] = 0'2158$

b) Si 2 coches aguardan cola es porque han llegado 12 coches (sólo puede atender a 10 coches/minuto)

$P[X=12] = \frac{(3'5)^{12}}{12!} e^{-3'5} = F(12)-F(11) = 0'9999 - 0'9997 = 0'0002$

$P[X=12] = 0'0002$

c) Si llegan más coches de los que puede atender en un minuto entrarán más de 10 coches

$P[X>10] = 1-P[X\leq 10] = 1-F(10) = 1-0'9990 = 0'0010$

d) ¿Qué número de coches debería ser capaz de atender por minuto esta estación de servicio si quisiera poder satisfacer la demanda en un minuto dado con probabilidad al menos de 0'95?

Nos piden un valor a tal que

$P[X\leq a] \geq 0'95 \Rightarrow F(a) \geq 0'95$. Si buscamos en las tablas: $a=7$ coches por minuto ya que

$P[X\leq 7] = F(7) = 0'9733 \geq 0'95$.

Ejemplo. Las erratas de imprenta de una página de cierto libro siguen una distribución de Poisson y hay 2 erratas de media por página. En un libro de 300 páginas: ¿cuál es la probabilidad de que en una o más páginas existan más de cinco erratas?

Hallamos la probabilidad de que una página contenga más de cinco erratas

X = nº de erratas en ... página

$P[X > 5] = 1 - P[X \leq 5] = 1 - F(5) = 1 - 0'9834 = 0'0166$

Sea Y: nº de páginas del libro que contienen más de cinco erratas. Entonces

$$Y \rightsquigarrow B(\underset{n}{300}, \underset{p}{0'0166})$$

Como n es muy grande frente a p

$B(300, 0'0166) \longrightarrow P(n \cdot p) = P(300 \cdot 0'0166) = P(4'98 \simeq 5)$

$B(300, 0'0166) \longrightarrow P(5)$. Nos piden:

$$P[Y \geqslant 1] = 1 - P[Y < 1] = 1 - P[y = 0] = 1 - \frac{5^0}{0!} e^{-5} = 0'9933$$

$P[Y \geqslant 1] = 0'9933$