ASTURIAS
Corporación Universitaria

Distribuciones Discretas

Índice

Objetivos

- Objetivo 1: Conocer los distintos tipos de distribuciones de tipo discreto.
- Objetivo 2: Comprender y analizar las diferentes distribuciones de variables discretas.

1 Distribución Uniforme Discreta

"Equiprobables: cuando hay *n* resultados diferentes y todos tienen la misma probabilidad de ocurrir"

Decimos que una variable aleatoria X tiene una **distribución uniforme discreta sobre el conjunto finito de números** $\{x_1,\dots,x_n\}$ **si la probabilidad de que** X **tome cualquiera de estos valores es la misma**, es decir, $1/n$. Esta distribución, **surge en espacios de probabilidad equiprobables**, esto es, en situaciones en donde **tenemos** n **resultados diferentes y todos ellos tienen la misma probabilidad de ocurrir**. Los juegos de lotería son un ejemplo donde puede aplicarse esta distribución de probabilidad. Escribimos entonces $X \sim unif\{x_1,x_2,\dots,x_n\}$ si

$$P(X=x)=\begin{cases}1/n & si\ x=x_1,x_2,\dots,x_n.\\ 0 & cualquier\ otro\ caso.\end{cases}$$

Ejemplo 1: La gráfica de la función de probabilidad de la distribución $unif\{1,2,3,4,5\}$ aparece en la Figura a continuación, junto con la correspondiente función de distribución. Cada salto de la función de distribución es de tamaño $1/5$. Es fácil ver que $E(X)=\frac{1}{n}\sum_{i=1}^{n}x_i$ y $Var(X)=\frac{1}{n}\sum_{i=1}^{n}(x_i-E(X))^2$.

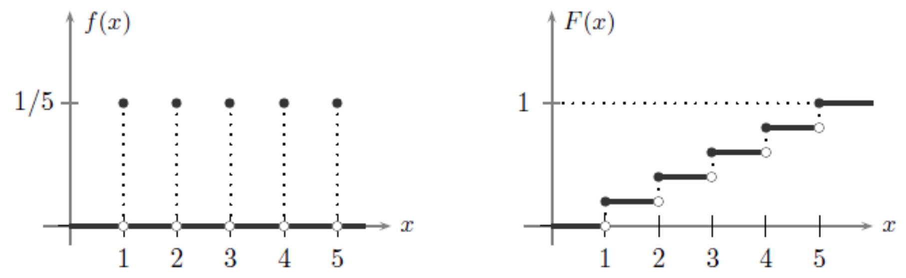


Figura 1: Distribución uniforme discreta"

2 Distribución Bernoulli

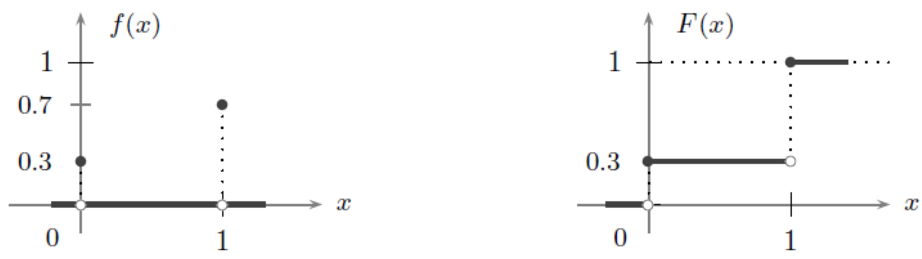
"Bernoulli: aquél experimento aleatorio con únicamente dos posibles resultados (éxito y fracaso)"

Un **ensayo Bernoulli** se define como **aquel experimento aleatorio con únicamente dos posibles resultados**, llamados genéricamente **éxito y fracaso**, con probabilidades respectivas p y $1-p$. Si se define la variable aleatoria X como aquella función que lleva el resultado **éxito** al número 1 y el resultado **fracaso** al número 0, entonces **decimos que** X **tiene una distribución Bernoulli con parámetro** $p\in(0,1)$, y escribimos $X\sim Ber(p)$. La función de probabilidad es entonces:

$$P(X=x)=\begin{cases}p^x(1-p)^{1-x} & Si\ x=0{,}1.\\ 0 & Cualquier\ otro\ caso.\end{cases}$$

La gráfica de la función de densidad de esta distribución para $p = 0.7$ aparece en la Figura mostrada a continuación, junto con la correspondiente función de distribución. En este caso, es muy sencillo verificar que $E(X) = p$ y $Var(X) = p(1-p)$. En la realización de todo experimento aleatorio, siempre es posible preguntarnos por la ocurrencia o no ocurrencia de un evento cualquiera. Por ejemplo: Ganar o no ganar en un juego de lotería, que llueva o no llueva hoy por la tarde, etc.
Este es el esquema general donde surge esta distribución, que aunque sencilla, es de amplia aplicación.

Figura 2: Distribución *Bernoulli*



3 Distribución Binomial

Supongamos ahora que tenemos una serie de *n* ensayos independientes *Bernoulli*, en donde la probabilidad de éxito en cualquiera de estos ensayos es p. Si denotamos por E el resultado éxito y por F el resultado fracaso, entonces el espacio muestral consiste de todas las posibles sucesiones de longitud n de caracteres E y F. Usando el principio multiplicativo, es fácil ver que el conjunto Ω tiene 2^n elementos. Si ahora definimos la variable aleatoria X como aquella que cuenta el número de éxitos en cada una de estas sucesiones, esto es, $X(EE \cdots EE) = n$, $X(FE \cdots EE) = -1$, . . ., $X(FF \cdots FF) = 0$, entonces tenemos que X puede tomar los valores $0, 1, 2, \ldots, n$ con las probabilidades que aparecen abajo. Decimos entonces que X tiene una *distribución binomial* con parámetros n y p, y escribimos $X \sim bin(n, p)$.

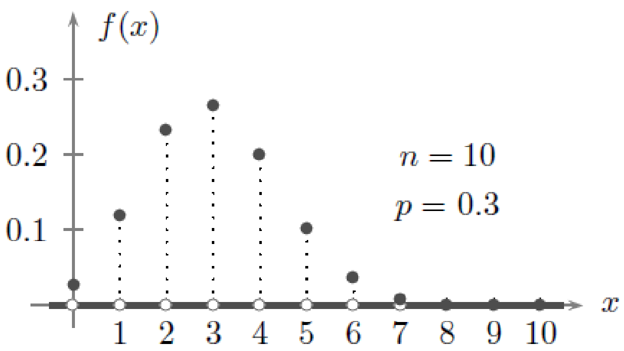
$$P(X = x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & Si\ x = 0,1,2, \ldots, n. \\ 0 & Culquier\ otro\ caso". \end{cases}$$

Ejemplo 2: Para $n = 10$ ensayos, con probabilidad $p = 0.3$, se puede calcular $P(X = 2) = \binom{10}{2}(0.3)^2(0.7)^{10-2}$ La gráfica de esta función de probabilidad con estos parámetros aparece en la Figura mostrada a continuación. La fórmula anterior puede justificarse de la forma siguiente. Queremos que en n ensayos *Bernoulli* se obtengan x éxitos y $n - x$ fracasos. La probabilidad de obtener esto es el número:

$$p \cdots p(1-p) \cdots (1-p) = p^x(1-p)^{n-x}$$

Pero hemos colocado los x éxitos en los primeros x ensayos, tenemos entonces que multiplicar por las diferentes formas en que estos x éxitos pueden distribuirse en los n ensayos, este factor es el coeficiente binomial $\binom{n}{x}$. Para esta distribución, puede demostrarse que $E(X) = np$, y $Var(X) = np(1-p)$.

Figura 3: Distribución *Binomial*



Ejemplo 3: En un conocido pueblo de la región ganadera, el 20% del ganado está afectado por una enfermedad A. Si se considera un grupo de 5 reses ¿Cuál es la probabilidad de que haya 2 reses enfermas?

Así, tenemos que, sea X el número de reses afectadas por A de entre 5. Entonces X tiene una distribución $bin(5, 0.2)$. Por tanto:

$$P(X=2) = \binom{5}{2}(0.2)^2(0.8)^3 = 0.2048$$

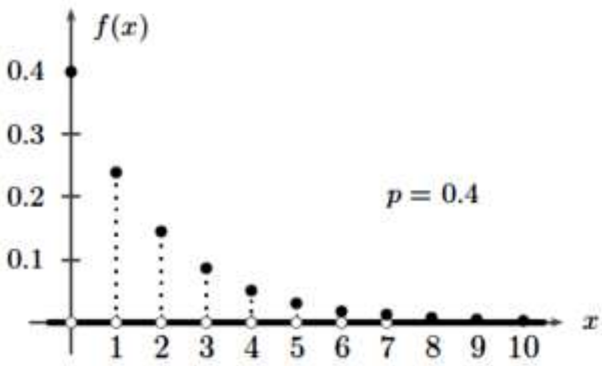
4 Distribución Geométrica

Supongamos que tenemos ahora una sucesión infinita de ensayos independientes *Bernoulli*, en cada uno de los cuales la probabilidad de éxito es p. Para cada una de estas sucesiones, definimos la variable aleatoria X como el número de fracasos antes de obtener el primer éxito. Por ejemplo, $X(FEFEFF\cdots) = 1$, $X(EFFEEE\cdots) = 0$, $X(FFFEFE\cdots) = 3$. Observamos que X puede tomar los valores $0, 1, 2, \ldots$ La probabilidad de que X tome el valor entero $x \geq 0$ es $p(1-p)x$. Decimos entonces que X tiene una distribución geométrica con parámetro p, y escribimos $X \sim geo(p)$ cuando:

$$P(X=x) = \begin{cases} p(1-p)^x & Si\ x = 0, 1, 2, \ldots \\ 0 & Cualquier\ otro\ caso". \end{cases}$$

Ejemplo 4: Para $p = 0.4$, la gráfica de esta función se muestra en la Figura a continuación. El nombre de esta distribución proviene del hecho de que cuando escribimos la suma de todas las probabilidades, obtenemos una suma geométrica. La inspección sucesiva de artículos hasta encontrar una defectuosa, posiblemente en un proceso de control de calidad, puede modelarse usando una distribución geométrica. Para esta distribución tenemos que $E(X) = (1 - p)/p$, y $Var(X) = (1 - p)/p2$.

Figura 4: Distribución *Geométrica*



En algunos textos se define la distribución geométrica contando el número de ensayos (No el de fracasos), antes del primer éxito. En este caso, la variable es $X + 1$, es decir, la distribución se desplaza hacia la derecha una unidad. La esperanza es ahora $1/p$ y la varianza permanece constante $(1 - p)/p^2$.

Ejemplo 5: En una cierta línea de ensamblaje de automóviles utilitarios, se sabe que en promedio uno de cada 150 automóviles ensamblados sale defectuoso. ¿Cuál es la probabilidad de que al hacer un control de calidad, el cuarto automóvil ensamblado sea el primero defectuoso?

Así, tenemos que, Sea D el suceso de seleccionar un automóvil defectuoso, sabemos que $P(D) = \frac{1}{150}$ dado que es un modelo de la distribución geométrica, la probabilidad pedida es:

$$P = \frac{1}{150}\left(\frac{149}{159}\right)^3 = 0.00653422$$

5 Distribución Poisson

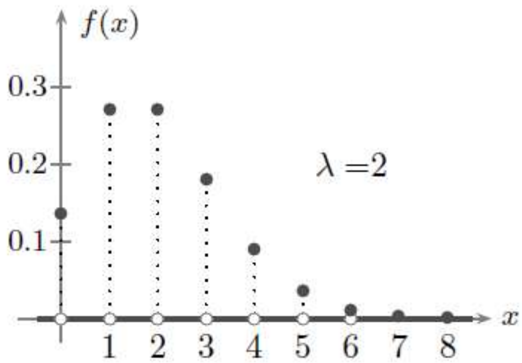
"Distribución Poisson: con ello observamos el número de ocurrencias de un cierto evento dentro de un intervalo de tiempo dado"

Supongamos que deseamos observar el número de ocurrencias de un cierto evento dentro de un intervalo de tiempo dado, por ejemplo, el número de clientes que llegan a un cajero automático durante la noche, o tal vez deseamos registrar el número de accidentes que ocurren en cierta avenida durante todo un día. Para modelar este tipo de situaciones podemos definir la variable aleatoria X como el número de ocurrencia de este evento en el intervalo de tiempo dado. Es claro entonces que X puede tomar los valores $0, 1, 2, \ldots$, y en principio no ponemos una cota superior para el número de observaciones del evento. Adicionalmente, supongamos que conocemos la tasa media de ocurrencia del evento de interés, que denotamos por la letra λ (lambda). El parámetro λ es positivo y se interpreta como el número promedio de ocurrencias del evento por unidad de tiempo. La probabilidad de que la variable aleatoria X tome un valor entero $x \geq 0$ se definirá a continuación. Decimos que X tiene una distribución *Poisson* con parámetro $\lambda > 0$, y escribimos $X \sim Poisson(\lambda)$ cuando:

$$P(X = x) = \begin{cases} e^{-\lambda}\dfrac{\lambda^x}{x!} & Si\ x = 0, 1\ , 2, \dots \\ 0 & Cualquier\ otro\ caso \end{cases}$$

Puede demostrarse que la función $f(x)$ arriba definida es efectivamente una función de probabilidad para cada valor de $\lambda > 0$. La forma de esta función se muestra a continuación cuando $\lambda = 2$. Después de algunos cálculos sencillo puede también comprobarse que $E(X) = \lambda$, y $Var(X) = \lambda$".

Figura 5: *Distribución Poisson. Tomada de (Rincón, 2010).p.56*



Puede además demostrarse que cuando $X \sim bin(n,p)$ y hacemos tender n a infinito y p a cero de tal forma que el producto np se mantenga constante igual a λ, entonces la variable aleatoria X adquiere la distribución *Poisson* con parámetro λ.

Este resultado sugiere que cuando n es grande, la distribución binomial puede ser aproximada mediante la distribución *Poisson* de parámetro $\lambda = np$. Esto es particularmente útil, pues el cálculo de probabilidades de la distribución *binomial* involucra el cálculo de factoriales y ello puede ser computacionalmente difícil.

Ejemplo 6: Una empresa de seguros contra fallos de equipos de computadora calculó que el 0.01% de los equipos tiene un cierto tipo de fallo cada año. ¿Cuál es la probabilidad de que esta empresa tenga que pagar por más de cuatro de sus 10000 computadores asegurados en un año dado?

Así, tenemos que, sea X el número de pagos efectuados en un año dado. X Distribución $bin(10000, 0.0001)$. Como $np = 1 < 5$ y

$p = 0.0001 < 0.1$, podemos aproximar la binomial por una Poisson $P(1)$. Entonces tenemos:

$$P(X > 4) = 1 - P(X \leq 4) = 1 - [P(X = 0) + \cdots + P(X = 4)] = 1 - \left[\frac{1^0}{0!}e^{-1} + \cdots + \frac{1^4}{4!}e^{-1}\right]$$
$$= 1 - [0.3679 + \cdots + 0.0153] = 0.0037.$$

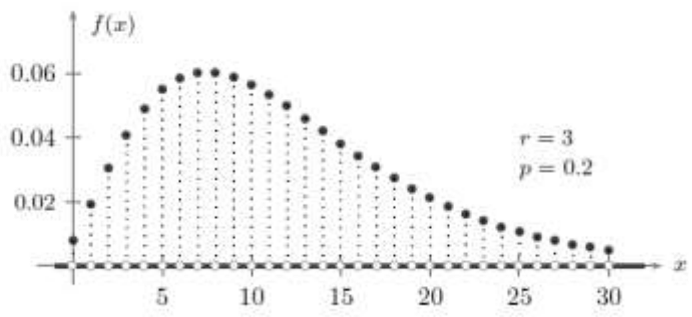
6 Distribución Binomial Negativa

Si en una sucesión infinita de ensayos *Bernoulli* la variable aleatoria X cuenta el número de fracasos antes de obtener el *r-ésimo* éxito, entonces decimos que X tiene una *distribución binomial negativa* con parámetros r y p, y escribimos $X \sim bin\,neg(r,p)$. En este caso, tenemos que X puede tomar los valores $0, 1, 2, \ldots$ con probabilidades como se indica a continuación:

$$P(X = x) = \begin{cases} \binom{r + x - 1}{x} p^r (1-p)^x & Si\ x = 0, 1, 2, \ldots \\ 0 & Cualquier\ otro\ caso \end{cases}$$

Aparece el término p^r pues la sucesión de ensayos Bernoulli no concluye sino hasta obtener r éxitos. Podemos tener un número variable de fracasos, de ahí el término $(1 - p)^x$, y finalmente el factor $\binom{r+x-1}{x}$ que nos dice las diferentes formas en que los r éxitos pueden aparecer en los r+x−1 ensayos realizados antes del último que necesariamente fue un éxito. La gráfica de esta función aparece en la Figura mostrada a continuación cuando $r = 3$ y $p = 0.2$.

Figura 6: *Distribución Binomial Negativa Tomada de* (Rincón, 2010).*p. 58*



Es claro, que esta distribución es una generalización de la distribución geométrica, la cual se obtiene tomando $r = 1$. Se puede además demostrar que $E(X) = r(1-p)/p$ y $Var(X) = r(1 - p)/p^2$.
La definición de coeficiente binomial puede extenderse del siguiente modo: Para cualquier número real a, y cualquier entero natural x se define

$\binom{a}{x} = a(a-1)\cdots(a-x+1)/x!$ Puede demostrarse la identidad

$\binom{n-x+1}{x} = (-1)^x\binom{-n}{x}$. De este hecho adquiere su nombre esta distribución.

7 Distribución Hipergeométrica

"Distribución hipergeométrica: conjunto de *N* objetos de los cuales *K* son de primera clase y *N-K* son de segunda clase"

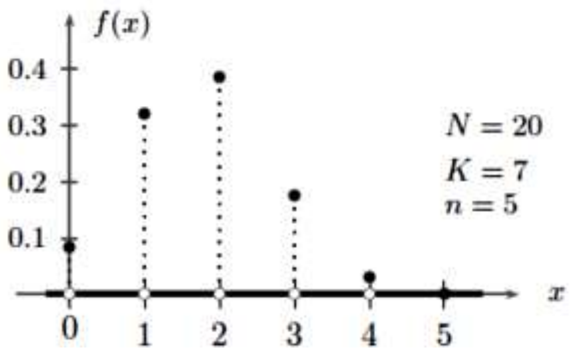
Supongamos que tenemos un conjunto de N objetos de los cuales K son de una primera clase y $N - K$ son de una segunda clase. Supongamos, que de este conjunto tomamos una muestra aleatoria de tamaño n, la muestra es sin reemplazo y el orden de los objetos seleccionados no importa. El espacio muestral de este experimento consiste, de todas las posibles muestras de tamaño n que se pueden obtener del conjunto

mayor de tamaño N. La cardinalidad del espacio muestral es $\binom{N}{n}$. Si para cada muestra definimos la variable aleatoria X como el número de objetos de la primera clase contenidos en la muestra seleccionada, entonces X puede tomar los valores $0, 1, 2, \ldots, n$, suponiendo $n \leq K$. La probabilidad de que X tome un valor x estará dada por la fórmula que enunciamos a continuación. Decimos que X tiene una distribución Hipergeométrica con parámetros N, K y n, y escribimos $X \sim hipergeo(N, K, n)$ si

$$P(X=x) = \begin{cases} \dfrac{\binom{K}{n}\binom{N-K}{n-x}}{\binom{N}{n}} & Si\ x = 0,1,\ldots,n \\ 0 & Cualquier\ otro\ caso \end{cases}$$

El término $\binom{K}{x}$ nos dice las diferentes formas en que de los K objetos de la primera clase se pueden escoger x de ellos, y el término $\binom{N-K}{n-x}$ es nuevamente las diferentes formas de escoger $n-x$ objetos de la totalidad de $N-K$ objetos de la segunda clase. Usamos el principio multiplicativo para obtener el número total de muestras diferentes en donde x objetos son de la primera clase y $n - x$ objetos son de la segunda clase. La gráfica de esta función de densidad para ciertos valores de los parámetros aparece en la Figura mostrada a continuación. En este caso es posible comprobar que $E(X) = nK/N$, y $Var(X) = n\frac{K}{N}\frac{N-K}{N}\frac{N-n}{N-1}$.

Figura 7: *Distribución Hipergeométrica*[1].



8 Conclusiones

- Equiprobables: cuando hay n resultados diferentes y todos tienen la misma probabilidad de ocurrir.
- Bernoulli: aquél experimento aleatorio con únicamente dos posibles resultados (éxito y fracaso).
- Distribución Poisson: con ello observamos el número de ocurrencias de un cierto evento dentro de un intervalo de tiempo dado.
- Distribución hipergeométrica: conjunto de N objetos de los cuales K son de primera clase y N-K son de segunda clase.

9 Referencias Bibliográficas

- Montgomery, D. C y Runger, R. (2008). Probabilidad y Estadística Aplicadas a la Ingeniería. 2da. Edición. Limusa Wiley. Mexico.
- Rincón. L. (2010). Curso Elemental de Probabilidad y Estadística. 1ra Edición. Circuito Exterior de CU. México D.F.
- Walpole, R. E., Myers, R. H., Myers,S. L. y Ye, K. (2007). Probabilidad y Estadística para Ingenieria y Ciencias. 8va. Edición. Pearson. Mexico
- I. Espejo Miranda, F. Fernández Palacín, M. A. López Sánchez, M. Muñoz Márquez, A. M. Rodríguez Chía, A. Sánchez Navas, C. Valero Franco. (2006). Estadística Descriptiva y Probabilidad. 3ra. Edición. Servicio de Publicaciones de la Universidad de Cádiz.
- Devore, J. (2008). Probabilidad y Estadística para Ingeniería y Ciencias. 7ma. Edición. Cengace Learning Editores. México

[1] Tomado de (Rincón, 2010).p.58

- Canavos, G. Probabilidad y Estadística. Métodos y Aplicaciones.
- Daniel, W. (2009). Bioestadística. Base para el análisis de las ciencias de la salud. Cuarta edición. Limusa Wiley. México.
- Mendenhall, W., Wackerly, D. y Scheaffer, R. (1994). Estadística Matemática con Aplicaciones. Grupo Editorial Iberoamericana. México.