UNIVERSIDAD TECNOLÓGICA NACIONAL

Facultad Regional San Nicolás

PROBABILIDAD y ESTADÍSTICA I

UNIDAD Nº3

Licenciatura en Enseñanza de la Matemática

Año 2011

Mg. Lucía C. Sacco

UNIDAD Nº3

Variables aleatorias unidimensionales. Distribuciones.

Noción general de una variable aleatoria. Variables aleatorias discretas y continuas. Definición y ejemplos.

Variable aleatoria discreta. Distribución de probabilidad para una variable aleatoria discreta. Función de distribución o función de probabilidad acumulada.
Características: Esperanza matemática y desvío estándar de una variable aleatoria discreta. Variancia y desvío estándar de una variable aleatoria discreta.

Distribuciones discreta: Bernoulli, Binomial, Hipergeométrica y de Poisson.

Distribución binomial de probabilidad. Aplicación. Características de la distribución binomial. Esperanza y desvío estándar de una distribución binomial.

La variable aleatoria de Poisson. Distribución de Poisson. Definición. Teorema. La distribución de Poisson como una aproximación a la distribución binomial.
Aplicación a problemas de espera. Ajuste de datos estadísticos a la distribución de Poisson.

Variables aleatorias continuas. La distribución normal. Propiedades. Distribución normal como límite de la distribución binomial. Integración de la función de probabilidad. Problemas de aplicación.

La ley de los grandes números. Teorema central del límite.

Propósitos:

Brindar oportunidades para la construcción de herramientas que permitan:

- Analizar los conceptos más relevantes de la teoría de probabilidades y aplicarlos a la resolución de problemas.
- Reconocer la necesidad del estudio de la teoría probabilística, como instrumento para medir la incertidumbre en el proceso inferencial y para la construcción de modelos que describan la realidad y posibiliten su análisis.
- Analizar diferentes distribuciones de probabilidad de variables aleatorias discretas y continuas con sus modelos matemáticos específicos.
- Distinguir campos de aplicación para cada modelo en particular tomando en cuenta especialmente la forma de selección de las unidades experimentales.

Bibliografía sugerida:

- Canavos, George. *Probabilidad y Estadística. Aplicaciones y Métodos*. México. McGraw Hill. 1988.
- Meyer Paul L. *Probabilidad y Aplicaciones Estadísticas*. México. Addison Wesley Iberoamericana .1993.
- Walpole Ronald, Myers Raymond. *Probabilidad y Estadística*. México. Pearson Educación. 1999.
- Apuntes Posgrado: "Estadística Aplicada a la Investigación". Facultad de Ciencias Económicas. Universidad Nacional de Córdoba.

1. Introducción

En la **Unidad Nº3** analizaremos distribuciones de probabilidad que corresponden a variables aleatorias discretas y continuas.

2. Variables aleatorias unidimensionales

Una variable aleatoria numérica es un fenómeno de interés cuyos resultados se pueden expresar en números.

Las variables aleatorias se clasifican en discretas y continuas:

- DISCRETA: La variable aleatoria X se dice que es discreta si los números asignados a los sucesos elementales de E son puntos aislados. Sus posibles valores constituyen un conjunto finito o infinito numerable. Por ejemplo, supongamos el experimento consistente en lanzar tres veces una moneda no trucada; si consideramos la variable aleatoria X=número de "caras" obtenidas en los tres lanzamientos, los valores que puede tomar esta variable aleatoria son finitos (0, 1, 2, 3).
- CONTINUA: La variable aleatoria X será continua si los valores asignados pueden ser cualesquiera, dentro de ciertos intervalos, es decir, puede tomar cualquier valor de ***R***. Por ejemplo, si consideramos el experimento aleatorio consistente en medir el nivel de agua en un embalse y tomamos la variable aleatoria X="nivel de agua", esta puede tomar valores entre 0 y más infinito.

Ejemplo 1

Supongamos un experimento aleatorio que consiste en tirar al aire simultáneamente dos monedas.

Los resultados posibles de dicho experimento son W = {cc, cx, xc, xx}

Si se está interesado en conocer el número de caras, al tirar dos monedas simultáneamente, necesitamos definir una variable que exprese cada resultado de este experimento aleatorio.

Entonces, simbolizando a esta variable con X, a la cantidad de caras al tirar dos monedas. De acuerdo al conjunto de resultados posibles de este experimento, podemos construir la siguiente tabla:

Cuantificación de Resultados de un experimento Aleatorio

Resultados del experimento	Valores de X
Cc	2
Cx	1
Xc	1
Xx	0

Los resultados de este experimento aleatorio no son números. Por eso, para poder cuantificarlos, tendremos que asociar un número a cada uno de estos resultados.

Una variable aleatoria es aquella cuyos valores surgen asignando números, a los resultados de un experimento aleatorio.

3. Distribución de probabilidad

Distribución de probabilidad es un modelo teórico que describe la forma en que varían los resultados de un experimento aleatorio, es decir, nos da todas las probabilidades de todos los posibles resultados que podrían obtenerse cuando se realiza un experimento aleatorio.

Se clasifican como discretas o continuas.

En la distribución de probabilidad discreta está permitido tomar sólo un número limitado de valores. En la continua, llamada función de densidad, la variable que se está considerando puede tomar cualquier valor dentro de un intervalo dado.

3.1 Distribución de Probabilidad Discreta

Sea un espacio probabilístico y sea X una variable aleatoria discreta que toma como posibles valores $x_1, x_2, \ldots x_n$.

Al conjunto de pares ordenados (x_i ; $P(x_i)$) se lo denomina **función de probabilidad** o **distribución de probabilidad** o simplemente distribución de la variable aleatoria x_i.

Lo simbolizamos en la siguiente tabla:

X	$P(X=x_i)$
x_1	$P(x_1)$
x_2	$P(x_2)$
x_3	$P(x_2)$
...	....

Las condiciones que cumple la **Distribución de Probabilidad** son:

(a) $P(x_i) \geq 0$ para cualquier valor de x_i

(b) $\sum_{i=1}^{n} P(x_i) = 1$ que se denomina **condición de cierre** de la distribución.

<u>Para el ejemplo 1:</u>

- simbolizaremos con x_i los valores particulares que toma la variable aleatoria X y en este caso se tendrá x1 = 0, x2 = 1, x3 = 2
- los resultados son igualmente probables y, por ello, es lógico pensar en asignarle a cada uno la misma probabilidad, o sea 1/4.

A partir de este razonamiento, construimos la siguiente tabla:

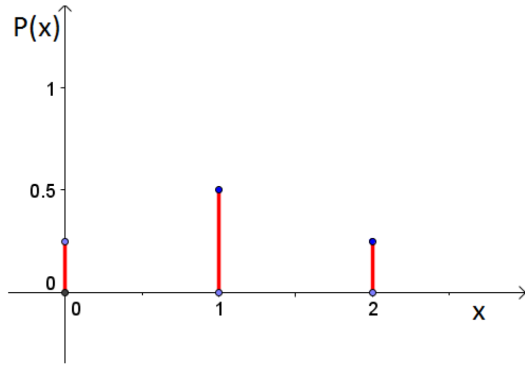
Función de Probabilidad de la Tirada Simultánea de Dos Monedas

Valores de la variable aleatoria	Probabilidad
0	1/4
1	2/4
2	1/4
	4/4 = 1

Se observa que en la tabla anterior, en la primera columna aparecen los distintos valores que toma la variable aleatoria X y en la segunda, la correspondiente probabilidad P(x) que puede asimilarse a la noción de frecuencia relativa.

Nótese que la suma de las probabilidades correspondientes a todos los valores posibles que toma la variable aleatorio es siempre igual a 1. Esto se debe a que los valores surgen de la cuantificación de todos los resultados posibles de un experimento aleatorio y ya hemos visto que P (S) = 1

Podemos representar gráficamente la función de probabilidad de una variable aleatoria discreta por medio de un gráfico a bastones.



En este caso, tendremos:

El **Valor esperado** de la variable es el promedio de todos sus valores, ponderados por sus respectivas probabilidades.

Ejemplo 2

Una ruleta contiene 38 casillas, numeradas 00, 0, 1, 2,.... 36. Una posibilidad consiste en apostar al suceso "el resultado es un número impar". Un jugador que realice esta apuesta ganará en 18 de los 38 resultados posibles.

Puesto que hay 18 impares:

P (El resultado es un número par) = 18/38 = 9/19

P (El resultado es un número impar) = 18/38 = 9/19

Si un jugador realiza una apuesta $1000, con lo cual recibe $2000 si el resultado es un número impar y nada en caso contrario, podemos representar la "ganancia" (en pesos) del jugador como $1000 con probabilidad 9/19 y (- $1000) con probabilidad (1-9/19) = 10/19. Entonces el valor esperado de su ganancia es:

1000 x (9/19) + (-1000) x 10/19 = - 100 pesos. Espera perder $100

3.2 Distribución de probabilidad continua

Si la variable aleatoria es continua, hay infinitos valores posibles de la variable y entre cada dos de ellos se podrían definir infinitos valores.

En estas condiciones no es posible deducir la probabilidad de un valor puntual de la variable como se puede hacer en el caso de las variables discretas. Pero sí es posible calcular la probabilidad acumulada hasta un cierto valor (función de distribución) y cómo cambia esa probabilidad acumulada en cada punto (densidad de probabilidad).

Por tanto, cuando la variable aleatoria sea continua hablaremos de **función de densidad**.

Sea X una variable aleatoria continua, se llama **función de densidad** y se representa como f(x) a una función no negativa definida sobre la recta real, tal que para cualquier x_i perteneciente al intervalo real $[a,b]$ se verifica $P(a \le X \le b) = \int_a^b f(x)\,dx$

Esta probabilidad se convierte en un área o superficie: bajo la curva f(x), sobre el eje $\overline{0x}$, entre las rectas de ecuación x = a y x = b.

4. Distribución de probabilidad para una variable aleatoria discreta

4.1 Función de Probabilidad

La función de distribución describe el comportamiento probabilístico de una variable aleatoria discreta X asociada a un experimento aleatorio y se representa como $F(x)$. Para estudiar la función de distribución distinguiremos entre el caso discreto y el caso continuo.

En este punto se analizará el caso discreto.

Sea X una variable aleatoria discreta asociada a un espacio probabilístico, se define la función de distribución:

$$F(x) : \Re \to [0,1] \text{ que verifica } F(x_i) = P[X \leq x_i] = \sum_{x_i < x} P_i$$

Estableciendo la analogía entre probabilidad y frecuencia relativa, resulta útil para el estudio de la distribución de la probabilidad de la variable discreta, obtener la función de distribución acumulada, la cual debe cumplir con las siguientes condiciones:

(a) $0 \leq F(x_i) \leq 1$

(b) Es una **función no decreciente.**
Dado dos números reales cualesquiera x_1 y x_2 tales que $x_1 < x_2$ entonces $F(x_1) < F(x_2)$

(c) $F(+\infty) = P(X \leq \infty) = 1$ ya que es pedir la probabilidad de que la variable toma cualquier valor de su campo de variación.

(d) $F(-\infty) = 0$ ya que no tomará valores negativos.

También es posible encontrar la $F(x_i)$ acumulada, la cual asigna a cada probabilidad la suma de las probabilidades acumuladas. Estos valores pueden ordenarse en una tabla como la siguiente:

x_i	$P(x_i)$	$F(x_i)$ acumulada
x_1	$P(x_1)$	$F(x_1)=P(x_1)$
x_2	$P(x_2)$	$P(x_1) + P(x_2) = F(x_2)$
....	....	....
....	....	....

Estas tablas pueden ser utilizadas para obtener diversas probabilidades para valores de x o para valores menores o iguales a x.

Ejemplo 3

Suponiendo el experimento aleatorio E: tirada de un dado y siendo la variable X: los puntos que aparecen en cada una de las tiradas.

La tabla de función de distribución de probabilidad acumulada:

x_i	$P(x_i)$	$F(x_i)$
1	1/6	1/6
2	1/6	1/3
3	1/6	1/2
4	1/6	2/3
5	1/6	5/6
6	1/6	1

Nos puede interesar calcular:

P (x = 3 puntos) = $\frac{1}{6}$; P (x ≤ 2 puntos) = $\frac{1}{3}$ y P (x > 6 puntos) = $\frac{0}{6} = 0$

Ejemplo 4

Sea E: "observar el resultado (Aprobado o no aprobado) de tres alumnos"

En este caso, los posibles resultados no vienen expresados en forma numérica. Se adopta entonces un código conveniente para expresar los resultados: Aprobó = 1 y No Aprobó = 0

De esta forma:

S = {(0,0,0) ; (0,0,1) ; (0,1,0) ; (1,0,0) ; (0,1,1) ; (1,0,1) ; (1,1,0) ; (1,1,1)}

Nos interesa estudiar el número de aprobados que se presentan y la denominaremos con x_i, cantidad variable que tomará valores diferentes cada vez que se realice el experimento. En nuestro ejemplo será:

$x_1 = 0$ aprobado {(0,0,0)}

$x_2 = 1$ aprobado {(0,0,1) ; (0,1,0) ; (1,0,0)}

$x_3 = 2$ aprobados {(1,1,0) ; (1,0,1) ; (0,1,1)}

$x_4 = 3$ aprobados {(1,1,1)}

En este caso la **variable aleatoria discreta** X = número de alumnos aprobados, ha sido generada por tres observaciones al azar y los valores posibles que pueden asumir son: $x_i = 0, 1, 2, 3$

Como cualquier valor de la variable aleatoria es un suceso, podemos calcular su probabilidad de aparición que denominaremos así:

$P(x_1 = 0) = \frac{1}{8}$ $P(x_2 = 1) = \frac{3}{8}$ $P(x_3 = 2) = \frac{3}{8}$ $P(x_4 = 3) = \frac{1}{8}$

Indicamos las funciones de probabilidad y la de distribución en la siguiente tabla:

x_i	$p(x_i)$	$F(x_i)$
0	$\frac{1}{8}$	$\frac{1}{8}$
1	$\frac{3}{8}$	$\frac{4}{8}$
2	$\frac{3}{8}$	$\frac{7}{8}$
3	$\frac{1}{8}$	$\frac{8}{8}$
	1	

Nos puede interesar calcular:

P (ningún aprobado) = $\frac{1}{8}$ P (todos aprobados) = $\frac{1}{8}$

P (al menos un aprobado) = $\frac{6}{8} = \frac{3}{4}$

Ejemplo 5

Se considera la siguiente distribución de frecuencias que corresponde a una distribución empírica de la variable edad en años de personas mordidas por perros.

Indicar la función de probabilidad y la función de distribución acumulada.

X: Edad de las personas mordidas por perros (en años)

Edad de las personas mordidas (años)	f_i	%	$P(x_i)$	$F(x_i)$
0 – 2	43	6,8	0,068	0,068
3 – 5	57	9,0	0,09	0,158
6 – 12	152	24,0	0,24	0,398
13 – 18	74	11,7	0,117	0,515
19 – 30	89	14,0	0,14	0,655
31 – 40	45	7,1	0,071	0,726
41 – 50	58	9,1	0,091	0,817
51 – 60	44	6,9	0,069	0,886
61 y más	72	11,4	0,114	1

Así, la probabilidad de que sea mordida una persona cuya edad oscila entre 3 y 5 años es 0.09.

Si queremos saber la probabilidad de que una persona menor de 18 años sea mordida debemos sumar:

$P(X \leq 18)$ = 0,068 + 0.09 + 0.24 + 0,117 = 0,515

Entonces el 51,5% de las personas mordidas tienen 18 años o menos.

4.2 Parámetros de una variable aleatoria discreta

Las medidas descriptivas para la distribución de una variable aleatoria se denominan parámetros.

Aprovechando la similitud entre frecuencia relativa y probabilidad, recordamos que cuando teníamos una distribución de frecuencias, calculábamos la media y la varianza.

Generalmente, cuando consideramos una variable aleatoria y su correspondiente función de probabilidad, la media aritmética de esta variable aleatoria se denomina **esperanza matemática.** La media μ de una distribución de probabilidad es el valor esperado o esperanza de su variable aleatoria.

La esperanza de una variable aleatoria discreta es un promedio ponderado de todos los resultados posibles.

La esperanza matemática de una variable aleatoria discreta, se calcula como la suma de cada valor que toma la variable multiplicado por su respectiva probabilidad

En símbolos $$E(X)=\sum_{i=1}^{n} x_i \; P(x_i)=\mu$$

Prácticamente, la esperanza matemática calcula el valor esperado promedio de una variable aleatoria el cual está en función de la probabilidad asignada a cada uno de los valores que toma dicha variable.

La varianza, de una variable aleatoria discreta, se define como la suma de los desvíos de cada valor que toma la Variable aleatoria con respecto a la esperanza matemática, elevados al cuadrado y multiplicados por su respectiva probabilidad.

En símbolos $$V(X)=\sum_{i=1}^{n}\left[x_i-E(X)\right]^2 P(x_i)$$

También en esta situación se puede obtener la desviación estándar como la raíz cuadrada positiva de la varianza.

$$\sigma=\sqrt{V(X)}=\sqrt{\sum_{i=1}^{n}\left[x_i-E(X)\right]^2 P(x_i)}$$

Un problema de aplicación

Supongamos que la probabilidad de que una casa de cierto tipo sea destruida por un incendio en un período cualquiera de 12 meses es 0,005, según cálculos actuariales efectuados por un organismo pertinente.

Una compañía de seguros ofrece al propietario de una casa una póliza contra incendio por el término de 1 año valuando la casa en 20.000 dólares y cobrándole una prima de 200 dólares.

¿Cuál es la ganancia esperada de la compañía?

Para resolver este problema debemos definir, en primer lugar, la variable aleatoria y establecer los distintos valores que ella puede tomar.

La variable aleatoria G = ganancia de la compañía, la cual puede tomar los valores 200 dólares si la casa no sufre un accidente de incendio y - 19.800 dólares si la casa se quema durante el año que cubre la póliza.

La función de probabilidad de G es, entonces:

G_1	$p(G_1)$
200	0,995
- 19.800	0,005
	1,000

Con esta información podemos calcular la ganancia esperada promedio de la compañía de seguros.

E (G) = 200 . 0,995 + (- 19.800) . 0,005 = 100 dólares

La compañía de seguros espera ganar 100 dólares, en promedio, con operaciones de este tipo.

Retomando el **ejemplo 3**, construimos la tabla para el cálculo de la esperanza y el desvío estándar:

x_i	$P(x_i)$	$x_i \; P(x_i)$	$[x_i - E(X)]^2$	$[x_i - E(X)]^2 \; P(x_i)$
0	1/8	0	2,25	0,28125
1	3/8	3/8	0,25	0,09375
2	3/8	6/8	0,25	0,09375
3	1/8	3/8	2,25	0,28125
	1	12/8=1,5		0,75

Es decir, el número esperado de alumnos aprobados es 1,5 (entre 1 y 2 alumnos), con un desvío estándar de 0.866. Dado que es un valor teórico, el resultado por más que corresponda a una variable aleatoria discreta no debe redondearse.

4.3 Distribuciones de probabilidad más conocidas

Las distribuciones discretas más conocidas en la literatura estadística son la distribución Bernoulli, la distribución Binomial, la distribución Hipergeométrica y la distribución de Poisson.

Definiremos brevemente cada una de ellas a través de las situaciones prácticas en las cuales se presentan y la manera de calcular sus respectivas probabilidades, ya sea por medio de su función de probabilidad y/o de tablas ya confeccionadas para tal fin.

4.3.1 Distribución Bernoulli

Según la definición de variable aleatoria discreta, ésta puede tomar sólo ciertos y determinados valores, generalmente expresados por números enteros.

El caso más sencillo es el que la variable aleatoria discreta es dicotómica, es decir que puede tomar sólo dos valores. Para simplificar, se puede pensar que estos dos únicos valores son el 0 y el 1.

Una variable aleatoria X que toma solamente los valores 0 y 1se denomina variable bipuntual o de Bernoulli.

Una variable bipuntual surge de un experimento llamado Bernoulli (en honor a Jacques Bernoulli que vivió en la última mitad del siglo XVII). Este experimento consiste en observar una sola unidad experimental y clasificarla en una de dos categorías mutuamente excluyentes y exhaustivas.

Ejemplo 6

Supongamos que se está estudiando el sexo de determinados individuos.

El experimento consiste en seleccionar un individuo y clasificarlo de acuerdo al sexo. Los dos únicos resultados posibles de este experimento serán W = (varón, mujer)

Para cuantificar estos dos resultados posibles y definir una variable aleatoria, podernos asignar el valor 1 al resultado varón y el valor 0 al resultado mujer. Estos dos resultados posibles del experimento de clasificar a un individuo según su sexo, serán igualmente probables. Es decir que puede asignarse una probabilidad igual a 1/2 a cada uno de estos eventos.

Sin embargo existen situaciones prácticas en las cuales los eventos se presentan con probabilidades diferentes. Si la muestra fuera extraída del conjunto de individuos que vive en una base militar, es evidente que el resultado varón tendría una mayor probabilidad de ocurrir.

Para poder asimilar esta distribución de probabilidad a cualquier situación, se simboliza la probabilidad de que el individuo sea varón con la letra p.

Si en cada repetición del experimento aleatorio se observa un solo individuo por vez:

- los resultados varón o mujer serán eventos mutuamente excluyentes. Como dos eventos mutuamente excluyentes son aquellos que, no pueden presentarse conjuntamente. Si el individuo que se observa es varón, no puede ser a la vez mujer.
- P (S) = 1, es decir, la probabilidad del conjunto definido por todos los resultados posibles de un experimento es igual a 1.

Si simbolizamos los dos eventos considerados en el ejemplo como V = varón y M = mujer, la probabilidad de que se observe un varón o una mujer es:

$$P\ (V \text{ o } M) = P(V \cup M) = P(V) + P(M) = P(W) = 1$$

Luego, si P (V) = p, entonces: P (M) = 1 - p

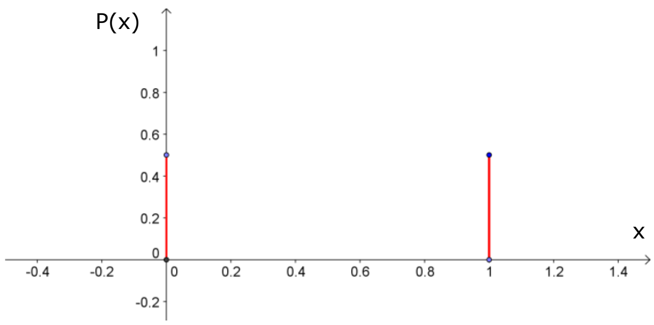
Una función de probabilidad de variable aleatoria discreta puede ser expresada por medio de una función matemática, una tabla o un gráfico.

A partir de estas consideraciones, podemos construir la siguiente tabla que representa la función de probabilidad de la variable aleatoria Bernoulli considerada:

Resultados del experimento	x	P(x)
V	1	p
M	0	1 - p
		1

De manera más general, la función puede expresarse por medio de una fórmula matemática $P(X = x) = p^x (1 - p)^{(1 - x)}$ para x = 0 ó x = 1

Esta función de probabilidad se grafica mediante el ya conocido gráfico de bastones; el que se utiliza siempre que consideramos una variable aleatoria discreta.



En nuestro ejemplo:

Donde P (varón) = p = 0,50 y P (mujer) = 1-p = 0,50

4.3.1.1 Parámetros de una Distribución Bernoulli:

- La esperanza matemática de una variable Bernoulli se obtiene sumando cada uno de los valores posibles de la variable, multiplicados por su probabilidad de ocurrencia: $E(X) = \sum_{i=0}^{1} x_i \; P(x_i)$

Reemplazando P (x_i) por P(X = x) = p^X $(1 - p)^{(1 - x)}$ obtenemos:

$$E(X) = \sum_{i=0}^{1} x_i \; p^{x_i} \; (1-p)^{1-x_i}$$

En tanto que, por definición, la variable aleatoria Bernoulli sólo puede tomar los valores 0 y 1, reemplazamos en la fórmula anterior xi en primer lugar por el valor 0 y luego por el valor 1 dentro de la Σ y tenemos:

$$E(X) = 0 \,.\, p^0 \,.\, (1 - p)^{1 - 0} + 1 \,.\, p^1 \,.\, (1 - p)^0 = 0 \,.\, 1 \,.\, (1 - p) + 1 \,.\, p \,.\, 1 = 0 + p = p$$

E(X) = p

La esperanza matemática de una variable con distribución Bernoulli es p.

- La varianza de una variable con distribución Bernoulli.

$$V(X) = \sum_{i=0}^{1} \left[x_i - E(X)\right]^2 P(x_i)$$

Volvemos a reemplazar x, primero por el valor 0 y luego por el valor 1, aplicando la fórmula de varianza ya conocida.

Para visualizar mejor el cálculo de la varianza es posible realizar la siguiente tabla:

x_i	$P(x_i)$	$[x_i - E(X)]^2$	$[x_i - E(X)]^2 \,.\, P(x_i)$
0	(1 – p)	$(0 - p)^2$	$(0 - p)^2 \,.\, (1 - p)$
1	P	$(1 - p)^2$	$(1 - p)^2 \,.\, p$

La varianza de X será la suma de la última columna de la tabla:

$$V(X) = (0 - p)^2 \,.\, (1 - p) + (1 - p)^2 \,.\, p = p^2\, (1 - p) + (1 - p)^2\, p$$

Ahora bien, sacando a (1 - p) como factor común, tenemos:

$$= (1 - p)\, [p^2 + (1 - p)\, p] = (1 - p)\, (p2 + p - p2) = (1 - p)\, p$$

Es común, en la literatura estadística llamar q a (1 - p) por lo que podemos escribir:

$$V(X) = p \,.\, (1 - p) = p \,.\, q$$

En consecuencia, la **desviación estándar** será:

$$D(X) = \sqrt{p \,.\, q}$$

Ejemplo 7

Se investiga una nueva vacuna y se estudia si, en un animal en particular, la inoculación es efectiva o no.

Si se ha establecido por ejemplo, que la vacuna es efectiva un 80% de las veces en que se aplica, p será igual a 0,80. Por lo tanto, la probabilidad de que un animal elegido sea sano será p = 0,80. Lógicamente, la probabilidad de seleccionar un animal enfermo será de 0,20.

Sano (x = 1) p = 0,80

Enfermo (x = 0) q = 0,20

4.3.2 Distribución Binomial

La distribución binomial es una distribución de probabilidad que permite describir fenómenos de diverso tipo biológicos, sociales, etc.

Para llegar a encontrar la ley binomial debemos estar en presencia de un experimento aleatorio en el que se realizan n representaciones de un experimento aleatorio que tiene las siguientes características:

1. Las observaciones posibles pueden obtenerse mediante 2 métodos de muestreo distinto. Cada observación puede ser seleccionada de una población infinita sin reemplazo o de una población finita con reemplazo[1].
2. Los posibles resultados son dos sucesos mutuamente excluyentes que designamos con A y $\bar{A}$.
3. El resultado de cada prueba es independiente de los resultados anteriores.
4. La probabilidad de que ocurra el suceso A se mantiene invariablemente de prueba en prueba y lo simbolizamos con *p*.
   Luego P (A) = p y P ($\bar{A}$) = 1 – *p* = *q*

La variable aleatoria considerada: *X* = número de veces que aparece A.

Ejemplo 8

Un biólogo que toma una planta y la clasifica de acuerdo a si presenta o no floración.

Suponiendo que este biólogo toma una muestra de 3 plantas y procede de la siguiente manera:

a) Toma la primera planta seleccionada en la muestra y la clasifica de acuerdo a sí tiene flores o no.

b) Repone esa planta, selecciona la segunda y la clasifica también de acuerdo a si presenta flores o no.

c) Procede de la misma manera con la tercera planta de la muestra.

Cada vez que el biólogo observa una planta seleccionada está realizando un experimento del tipo Bernoulli[2]. Recapitulando:

a) Se repite varias veces un experimento del tipo Bernoulli siendo sus resultados independientes.

b) En cada prueba, el resultado puede clasificarse en una de dos categorías mutuamente excluyentes que denominamos "éxito" o "fracaso".

c) La probabilidad p de éxito en cada prueba (que en este caso es encontrar una planta con flores), se mantiene constante de prueba a prueba.

d) El experimento se realiza un número n predeterminado de veces.

[1] Ver "Muestreo con remplazo y muestreo sin reemplazo" en el anexo al final de esta unidad.
[2] Un experimento de tipo Bernoulli surge cuando se selecciona al azar una unidad experimental y, en base a características predeterminadas, se la clasifica en una de dos categorías mutuamente excluyentes y exhaustivas.

En general, cuando el experimento presenta las características enunciadas precedentemente y la variable aleatoria de interés se define como "la cantidad de veces que se presenta uno de los dos resultados posibles, generalmente denominado "éxito", se tiene una variable con distribución binomial".

Una variable con distribución binomial se define como la cantidad de veces que se presenta uno de los dos resultados posibles, generalmente denominado éxito, en una serie finita de repeticiones de experimentos del tipo Bernoulli.

Volviendo al ejemplo del biólogo, los resultados posibles del experimento que consistía en observar la floración de plantas son los siguientes:

Sea F = la planta tiene flores (éxito) y $\overline{F}$ = la planta no tiene flores (fracaso)

Definimos la variable aleatoria X = cantidad de plantas con flores = cantidad de éxitos en las n = 3 repeticiones del experimento.

Considerando P (F) = p y P ($\overline{F}$) = 1 – p = q

El primer resultado de la tabla (FFF) indica que la primera planta observada tiene flores y la segunda tiene flores y la tercera tiene flores.

Asimilando este resultado a la teoría de probabilidades, podemos simbolizarlo como:

$$P\ (F \text{ y } F \text{ y } F) = P\ (F \cap F \cap F)$$

Al ser definida cada repetición del experimento como independiente de las demás, se tiene:

$$P\ (F \cap F \cap F) = P\ (F) \,.\, P\ (F) \,.\, P\ (F) = p \,.\, p \,.\, p = p^3$$

(Ya que establecimos que p es la probabilidad de éxito - planta con flores).

Como q expresa la probabilidad de fracaso, de la misma manera se establecieron las probabilidades de los demás resultados del experimento aleatorio.

Podemos así construir la siguiente tabla:

Eventos	X	Probabilidad
FFF	3	$ppp = p^3$
$FF\overline{F}$	2	$ppq = p^2q$
$F\overline{F}F$	2	$pqp = p^2q$
$\overline{F}FF$	2	$qpp = p^2q$
$F\overline{F}\,\overline{F}$	1	$pqq = pq^2$
$\overline{F}F\overline{F}$	1	$qpq = pq^2$
$\overline{F}\,\overline{F}F$	1	$qqp = pq^2$
$\overline{F}\,\overline{F}\,\overline{F}$	0	$qqq = q^3$

A partir de ella y reacomodando convenientemente las probabilidades, obtenemos:

x	P(x)
3	p^3
2	$3\,.\,p^2.q$
1	$3\,.\,p.q^2$
0	q^3
	1

Puede verificarse que la suma de P(x) = 1 a partir del desarrollo del cubo de un binomio $p^3 + 3p^2.q + 3p.q^2 + q^3 = (p + q)^3$. Y como p + q = 1, luego $(p + q)^3 = 1$.

Hemos expresado la función de probabilidad de una variable binomial por medio de una tabla. De ella podemos deducir la generalización de la función de probabilidad de una variable binomial, es decir $P(X = x) = C_n^x\, p^x\,.\,q^{n-x}$

Esta fórmula matemática expresa la función de probabilidad de una variable aleatoria discreta binomial. Representa la distribución de probabilidad binomial para obtener un número de éxitos x, si se conocen los parámetros de la distribución n y p.

La estructura de la fórmula binomial consta de dos partes:

- una de ellas corresponde al término $p^x \ q^{n\ -\ x}$. Este nos indica que cualquier resultado del experimento debe tener x éxitos y n - x fracasos en el conjunto de n pruebas.
- la otra parte expresa la cantidad de maneras mutuamente excluyentes en las que los x éxitos y los n - x fracasos pueden ser ordenados. En álgebra, a este término, se lo denomina número combinatorio o combinaciones de n elementos tomados de x en x, que es igual a $C_n^x = \dfrac{n!}{x!\,(n-x)!}$

Si en el ejemplo del biólogo nos interesa encontrar la probabilidad de obtener 2 plantas con flores al seleccionar 3 plantas, o sea, P(X = 2), esta probabilidad tiene 3 resultados favorables de los 8 resultados posibles, que son:

$$F\ F\ \overline{F} \qquad F\ \overline{F}\ F \qquad \overline{F}\ F\ F$$

Hay tantos resultados favorables al evento "obtener 2 plantas con flores", como la forma de seleccionar 2 objetos a un mismo tiempo a partir de un conjunto de 3 objetos.

Esta cantidad se puede obtener utilizando el número combinatorio $C_3^2 = \dfrac{3!}{2!\ 1!}$ que se lee combinaciones de 3 elementos tomados de 2 en 2, obteniendo $C_3^2 = \dfrac{3.2.1}{2.1\,1} = 3$

Esto significa que hay tres maneras de elegir dos plantas con flores, en un conjunto de tres plantas. Hay tres resultados posibles a este evento.

P ($F\ F\ \overline{F} \cup F\ \overline{F}\ F \cup \overline{F}\ F\ F$) = P ($F\ F\ \overline{F}$) + P ($F\ \overline{F}\ F$) + P ($\overline{F}\ F\ F$) =

= p^2 q + p^2 q + p^2 q = C_3^2 p^2 q =3 p^2 q

La probabilidad de la unión de estos eventos se calcula como la suma de las probabilidades, ya que los 3 son eventos mutuamente excluyentes.

Ahora bien, si el experimento se realiza un número n finito de veces, los n + 1 posibles valores de x serán: x = 0, 1, 2, ..., n.

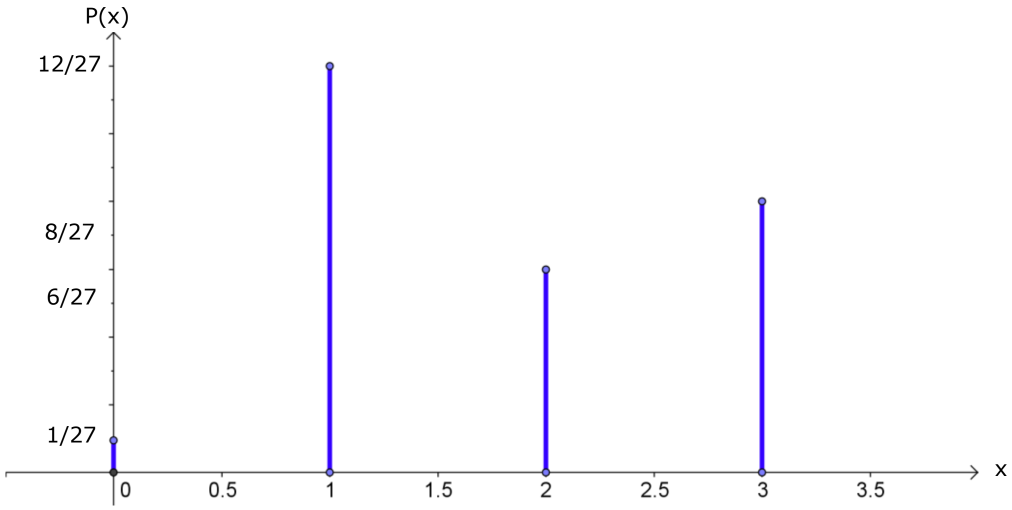
Ya vimos en el ejemplo de las plantas que, al observar n = 3 plantas, X = cantidad de plantas con flores, puede tomar los valores 0, 1, 2 y 3.

Supongamos que la probabilidad de encontrar una planta con flores en este experimento en particular es p = 2/3. La probabilidad de encontrar una sin flores será 1 - p = q = 1/3.

En este caso construimos la siguiente tabla que expresa la función de probabilidad como:

x	p(x)
0	$C_3^0\,(2/3)^0 \,.\, (1/3)^3 = 1 \,.\, 1 \,.\, \dfrac{1}{27} = \dfrac{1}{27}$
1	$C_3^1\,(2/3)^1 \,.\, (1/3)^2 = 3 \,.\, \dfrac{4}{9} \,.\, \dfrac{1}{3} = \dfrac{4}{9}$
2	$C_3^2\,(2/3)^2 \,.\, (1/3)^1 = 3 \,.\, \dfrac{1}{27} \,.\, \dfrac{1}{9} = \dfrac{2}{9}$
3	$C_3^3\,(2/3)^3 \,.\, (1/3)^0 = 1 \,.\, \dfrac{8}{27} \,.\, 1 = \dfrac{8}{27}$

La representación gráfica de la función de probabilidad de esta variable binomial es:



Como estamos considerando todos los resultados posibles de un experimento Bernoulli, la suma de todas las probabilidades asignadas debe ser igual a 1.

Es posible demostrar esta afirmación el teorema binomial, de la siguiente manera:

$$\sum_{x=0}^{n} P(X=x) = \sum_{x=0}^{n} C_n^x \, p^x \, . \, q^{n-x} = (p+q)^n = 1^n = 1$$

4.3.2.1 Parámetros de una Distribución Binomial:

La variable binomial surge de la repetición de experimentos Bernoulli; por eso, la esperanza matemática se calcula sumando n veces la esperanza matemática de la variable bipuntual que es p, o sea:

$$E(X) = (p + p + p ... + p) = n \, . \, p \qquad donde \quad p = \frac{M}{N}$$

Siendo M la cantidad de elementos de la población que pertenecen a la categoría definida como éxito y N el total de elementos de la población.

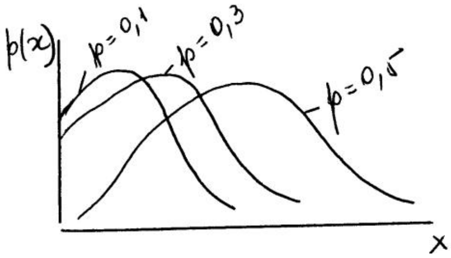
- ***La esperanza matemática de una variable con distribución binomial es*** $E(X) = n \, . \, p$
- ***La varianza de una variable con distribución binomial es*** $V(X) = n \, . \, p \, . q$
- ***La desviación estándar de una variable con distribución binomial es la raíz cuadrada positiva de*** $D(X) = \sqrt{n \, . \, p \, . q}$

Debido a la complejidad matemática para calcular probabilidades de variables binomiales, existen tablas donde se pueden encontrar estas probabilidades ya calculadas. Consultar documento que aparece en:

http://www-eio.upc.es/teaching/estad/MC/taules/com-usar-taules.pdf

4.3.2.2 Características de una Distribución Binomial:

Cada vez que se indican los parámetros n y p, se origina una distribución de probabilidad binomial. La forma de la distribución depende de *n* y de *p* que son los parámetros. Observamos a continuación varias distribuciones binomiales para distintos valores de p.



A medida que *p* se acerca a 0,5 la distribución se hace más simétrica y más grande la dispersión. Al aumentar *n*, para cualquier valor de *p*, la asimetría es menos pronunciada.

4.3.3 Distribución Hipergeométrica

La distribución hipergeométrica es una distribución cercanamente relacionada con la distribución binomial.

La situación experimental en la que se define una variable hipergeométrica también presume la extracción de elementos de una determinada población a los que podemos clasificar en una de dos categorías mutuamente excluyentes y exhaustivas denominadas éxito y fracaso.

En la práctica, la mayoría de los muestreos se efectúan sin reemplazo, ya que seleccionar al individuo varias veces en la muestra, no tiene sentido. Pero cuando la muestra es muy pequeña en relación al tamaño de la población de la que proviene, debemos considerar que estamos en presencia de un esquema binomial.

Veamos dos ejemplos para analizar si el muestreo debe hacerse con o sin reemplazo.

Problema de aplicación (a)

Supongamos, por ejemplo, que se debe extraer una muestra de tamaño n = 100 de una población de 1.000.000 de habitantes donde se conoce que 600.000 son varones. Veremos entonces qué pasa al extraer individuos de esta población.

La probabilidad de que el primer individuo extraído sea varón será:

600.000/1.000.000 = 0,60. La probabilidad de extraer el segundo individuo varón será: 599.999/999.999 = 0,5999, la de extraer el tercero será:

599.998/999.998 = 0,5999 y así sucesivamente. Si comparamos estos cocientes, vemos que la probabilidad prácticamente no varía de repetición en repetición.

En este caso particular podemos adaptar un modelo de muestreo con reemplazo.

Problema de aplicación (b)

En la reunión de una cooperadora escolar a la que asisten 20 personas, entre las que hay 14 mujeres, se quiere seleccionar una muestra de 5 personas para integrar una subcomisión.

Si definimos una variable aleatoria X como la cantidad de mujeres que integran la comisión, vemos que la probabilidad de extraer la primera mujer será 14/20 = 0,70; la probabilidad de extraer la segunda será de 13/19 =0,684; la de extraer la tercera de 12/18 = 0,67 y así sucesivamente.

Vemos aquí, que la probabilidad condicionada de extraer mujeres de esta población, no se mantiene constante en cada repetición del experimento.

Por este motivo, no es posible aplicar el esquema binomial. Estamos ante la situación de un muestreo sin reemplazo donde la probabilidad de extraer un elemento de la población varía en cada repetición.

En este caso, la variable así definida sigue una distribución denominada hipergeométrica.

La distribución hipergeométrica se aplica al muestreo sin reposición de una población finita, cuyos elementos pueden ser clasificados en dos categorías.

Antes de expresar la fórmula correspondiente a una función de probabilidad hipergeométrica, tendremos en cuenta la siguiente simbología:

N = cantidad de elementos de la población.

M = cantidad de elementos de la población que pertenecen a la categoría éxito.

N - M = cantidad de elementos de la población que pertenecen a la categoría fracaso.

n = cantidad de elementos en la muestra seleccionados de los N elementos de la población.

x = cantidad de éxitos en la muestra.

n - x = cantidad de elementos de la muestra que pertenecen a la categoría fracaso.

La función de probabilidad correspondiente a una variable aleatoria hipergeométrica, es la siguiente:

$$P(X = x) = \frac{C_M^x \cdot C_{N-M}^{n-x}}{C_N^n}$$

En síntesis:

Una variable aleatoria con distribución hipergeométrica se define como la cantidad de individuos que pertenecen a una cierta categoría, extraídos de una población dicotómica por medio de un muestreo sin reemplazo.

4.3.3.1 Parámetros de una Distribución Hipergeométrica:

Dejando de lado las operaciones involucradas, diremos que la expresión de la esperanza matemática es:

$$E(X) = \frac{n\ M}{N} = n\ p \qquad donde \quad p = \frac{M}{N}$$

Donde M/N es la proporción de éxitos en la población (similar al p de la distribución binomial).

A su vez, la varianza de una variable hipergeométrica es:

$$V(X) = \frac{n\,M}{N} \cdot \frac{(N-M)}{N} \cdot \frac{N-n}{N-1} = n\ p\ q . \frac{(N-n)}{(N-1)}$$

La esperanza matemática es la misma que la vista para la distribución binomial. La varianza también es la misma, pero multiplicada por el factor (N - n)/(N - 1), conocido en estadística como **factor de corrección para poblaciones finitas**.

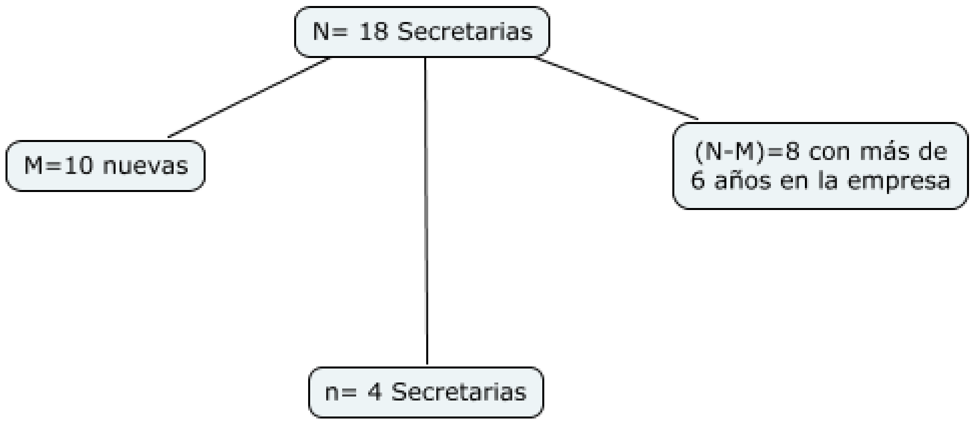
El cálculo de probabilidades hipergeométricas es muy complicado si no se tiene un programa de computación adecuado, ya que se deben calcular factoriales de alto orden.

Ejemplo 9

Un importante estudio contable tiene 18 secretarias; 8 de las cuales han estado en la firma por más de 6 años. Si un ejecutivo desea seleccionar 4 secretarias para asignarles una tarea nueva:

¿Cuál es la probabilidad de que en dicha comisión haya 2 de ellas que tengan menos de 6 años de experiencia?

De acuerdo a la información dada es posible organizar la información:



Si se define la variable aleatoria X como la cantidad de éxitos en n pruebas, será $X = 2$, ya que en este caso, la categoría éxito corresponde a la extracción de secretarias con menos de 6 años en la firma.

- La probabilidad hipergeométrica:

$$P(X = 2) = \frac{C_{10}^{2} . C_{8}^{2}}{C_{18}^{4}} = \frac{10!\, 8!\, 4!\, 14!}{2!\, 8!\, 2!\, 6!\, 18!} = 0{,}4118$$

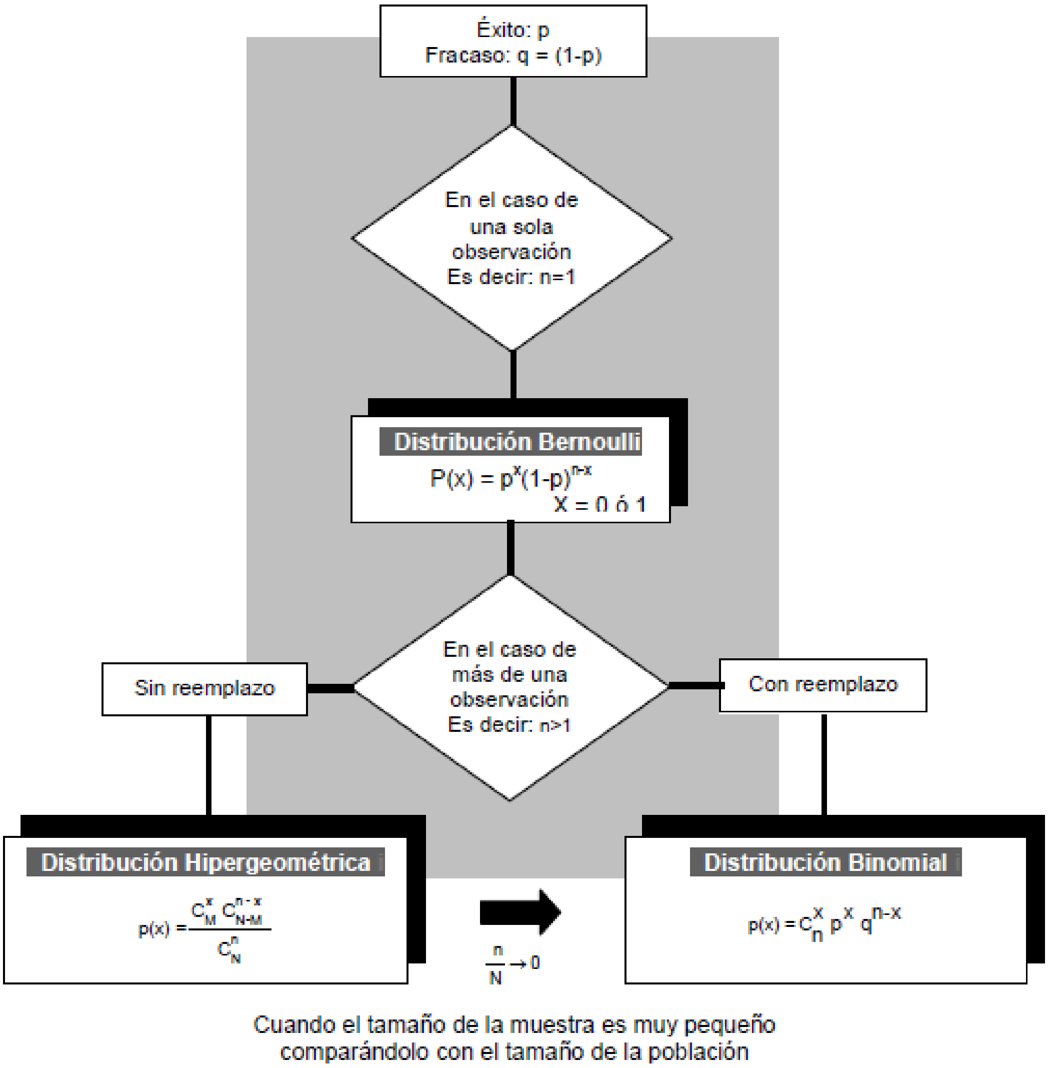
- El número esperado de secretarias nuevas:

$$E(X) = \frac{4\,.\,10}{18} = 2{,}22 \quad \text{Secretarias nuevas}$$

- La varianza de esta distribución será:

$$V(X) = 4\,.\,\frac{10}{18}\,.\,\frac{8}{18} . \frac{(18-4)}{(18-1)} = 0{,}8134$$

Podemos resumir hasta ahora lo desarrollado entre las tres distribuciones desarrolladas, Bernoulli, Binomial e Hipergeométrica.



4.3.4 Distribución de Poisson

La distribución de Poisson puede ser estudiada desde dos puntos de vista, como distribución límite de la distribución binomial o como proceso estocástico.

Trataremos de estudiarla desde estos dos puntos de vista de la manera más sencilla posible.

4.3.4.1 Distribución de Poisson como límite de la distribución binomial

Para estudiar la distribución binomial, siempre se consideran muestras de tamaño pequeño. En ellas, las unidades experimentales podían pertenecer a una de dos categorías (atacados o no atacados por hepatitis B, insectos vivos o muertos, supermercados que compran o no compran, etc.).

Pero en muchos casos nos encontraremos ante muestras de tamaño bastante mayor y si, además, la probabilidad de que ocurra lo que hemos denominado un éxito es pequeña, en lugar de calcular probabilidades por medio de la distribución binomial, utilizaremos una nueva distribución de variables discretas conocida como distribución de Poisson.

Ella se ocupa del número de veces que ocurre un acontecimiento raro ("p" chico), ya que a diferencia de la distribución binomial, es muy grande el número de veces que el acontecimiento no ocurre.

En general se toma por convención m = np < 5; es decir np menor a cinco unidades.

La función de probabilidad de una variable poissoniana X está dada por:

$$P(X = x) = \frac{e^{-m}.m^{x}}{x!}$$

Donde m > 0, x = 0, 1, 2, ..., y e: número neperiano (constante matemática).

En la distribución de Poisson el parámetro m es tanto la esperanza matemática como la varianza de la distribución.

$$E(X) = V(X) = m$$

Ejemplo 10

Supongamos que un agrónomo sabe, a través de la literatura existente sobre el tema, que el 2% de las semillas correspondientes a una especie determinada, no germinan en un tipo específico de suelo.

A los fines de tomar los recaudos pretinentes, antes de comenzar cierta investigación, desea conocer la probabilidad de que, en el caso de sembrar 100 semillas, 6 de ellas no germinen.

En primer lugar, debemos calcular el valor del parámetro "m", esperanza matemática de la distribución.

Entonces, como estamos considerando la distribución de Poisson como límite de la distribución binomial, la esperanza matemática se calcula igual que en este caso.

Teniendo en cuenta que m=n.p con n=cantidad de semillas en la muestra y p= proporción de semillas que no germinan, obtenemos:

m = np = 100 **.** 0,02 = 2 semillas no germinadas

Luego, definiendo a la variable en estudio X = cantidad de semillas no germinadas en la muestra de 100 semillas, la probabilidad resulta:

$$P(X = 6) = \frac{e^{-2}.2^{6}}{6!} = 0{,}012$$

Existen tablas donde se encuentran ya calculadas las probabilidades correspondientes a una variable con distribución de Poisson.

Consultar documento que aparece en:

http://www-eio.upc.es/teaching/estad/MC/taules/com-usar-taules.pdf

Buscando convenientemente en la tabla, obtenemos P(X = 6) = 0,0120. Este valor nos indica que existe una probabilidad de 0,012 de que no germinen 6 semillas en la muestra de 100 semillas.

Como vemos en este ejemplo, el tamaño de la muestra es grande, n = 100 y la probabilidad de éxito, que en este caso es la no germinación de una semilla, es muy baja ya que p = 0,02.

En realidad, el esquema del experimento coincide con el expuesto en el caso de una variable con distribución binomial, a saber:

Una unidad experimental (la semilla) puede ser asignada a una de dos categorías (germina o no germina). A su vez, la variable en estudio se define como la cantidad de veces que se da una categoría en particular que se denomina éxito. En este caso, la variable aleatoria es la cantidad de semillas que no germinan.

Lo que hemos hecho es asimilar esta distribución binomial en origen, a una distribución Poisson, ya que el tamaño de la muestra es grande y la probabilidad de éxito es pequeña.

4.3.4.2 Distribución de Poisson como proceso estocástico

Otra manera de definir la distribución de Poisson es considerando que surge a través de un proceso aleatorio.

Esta consideración establece una variable poissoniana distribuyéndola en forma espacial o temporal a lo largo de una muestra. A su vez, esta distribución debe ser estrictamente al azar o aleatoria o estocástica.

Ejemplo 11

Un investigador en ecología está estudiando la distribución de plantas de musgos en la ladera de una montaña.

Para eso, divide la superficie del terreno en estudio en cuadrículas del mismo tamaño marcadas con un hilo. Dentro de cada cuadrícula así establecida, cuenta la cantidad de plantas de musgo. Lo que está haciendo este investigador, precisamente, estudiar la distribución espacial de las plantas de musgo en el terreno que eligió para hacer la investigación. Además, no existen razones para suponer que las plantas de musgo no se distribuyen aleatoriamente a través de las cuadrículas, si ellas fueron marcadas objetivamente.

Gráficamente, tendríamos:

XX	X		XXX	X
X	XX	XXX	X	XX
	X	XX	X	X

La variable de una distribución de Poisson X se define ahora como la cantidad de eventos que ocurren por unidad de tiempo o de espacio y toma valores desde 0 en adelante.

Una variable poissoniana definida a través de un proceso estocástico debe cumplir las siguientes propiedades:

1) Su correspondiente media o esperanza matemática debe ser pequeña en comparación con el número máximo de eventos ocurridos en cada unidad de tiempo o de espacio. En otras palabras, el evento debe ser 'raro'. En el ejemplo dado en primera instancia, los cuadrados donde se cuentan las plantas de musgo deben ser lo suficientemente grandes como para que las plantas de musgo puedan crecer y desarrollarse físicamente. Así, un cuadrado de 1 cm. de lado es muy pequeño para que pueda desarrollarse una gran cantidad de plantas de musgo y considerar su distribución espacial según una distribución de Poisson.

   En cuanto a los fenómenos distribuidos temporalmente, la longitud del intervalo de tiempo debe ser lo suficientemente grande como para que, la ocurrencia del evento pueda seguir una ley de comportamiento poissoniana (ser raros).

2) La aparición de un éxito debe ser independiente de los éxitos anteriores en una misma unidad de tiempo o de espacio.

De esta manera, la aparición de una planta de musgo en una cuadrícula del campo no debe impedir ni favorecer la presencia o el desarrollo de otras plantas de musgo.

En síntesis, los eventos que poseen las propiedades de raros y aleatorios se distribuyen mediante una distribución de Poisson.

Ejemplo 12

Se estudia la distribución de células de levadura en 400 cuadrículas de un hemacitómetro. Este es un aparato semejante al utilizado para hacer los recuentos de las células sanguíneas y de otras estructuras microscópicas suspendidas en líquidos.

Después de haber muestreado 400 cuadrículas y contado la cantidad de células de levadura en cada una de ellas, se obtuvo la siguiente tabla de distribución de frecuencias:

Cantidad de células por cuadrícula	Cantidad de cuadrículas observadas
0	75
1	103
2	121
3	54
4	30
5	13
6	2
7	1
8	0
9	1
	400

Puede observarse que hay 75 cuadrículas que no presentan ninguna célula de levadura y que la mayor parte de las cuadrículas poseen sólo 1 o 2 células.

Por el contrario, únicamente 17 cuadrículas contienen 5 o más células de levadura.

¿Qué nos hace pensar que esta distribución de frecuencias siga una ley de Poisson?

Empecemos calculando la esperanza matemática o promedio de esta distribución:

$$\bar{x} = 1{,}8 \text{ células por cuadrícula}$$

Vemos que se trata de un valor muy pequeño, considerando que la cantidad de células que podrían existir en una cuadrícula es muy grande.

Esto nos lleva a pensar que estamos en presencia de un acontecimiento relativamente raro. También podríamos esperar que la existencia de células de levadura en un cuadrado, fuese independiente de la existencia de este tipo de células en otros cuadrados.

Así, la observación de los resultados de este experimento cumple con las propiedades que hemos enunciado al considerar una distribución de Poisson como proceso estocástico.

Ahora bien, si quisiéramos calcular la probabilidad de que, por ejemplo, se observen 2 células de levadura en una cuadrícula en particular, debemos considerar el único parámetro que caracteriza a una distribución de Poisson: "m". En este ejemplo:

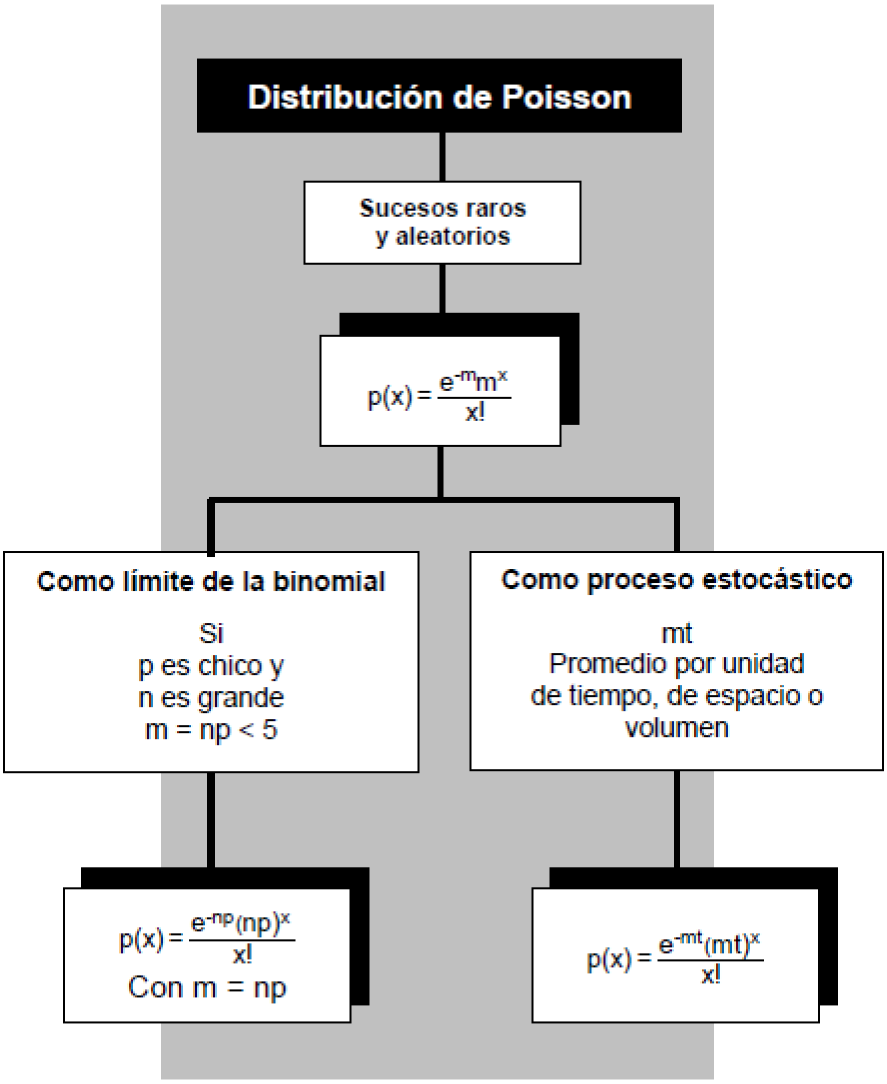
$$m = 1{,}8.$$

Luego:

$$P(X = 2) = 0{,}2678$$

Según lo obtenido utilizando la tabla de Poisson.

SÍNTESIS DE LA DISTRIBUCIÓN DE POISSON



La distribución de Poisson es importante debido a los numerosos fenómenos aleatorios que parecen seguirla.

5. Distribución de probabilidad de variables aleatorias Continuas

Veremos, ahora, como se comportan las funciones de probabilidad cuando la variable aleatoria es continua.

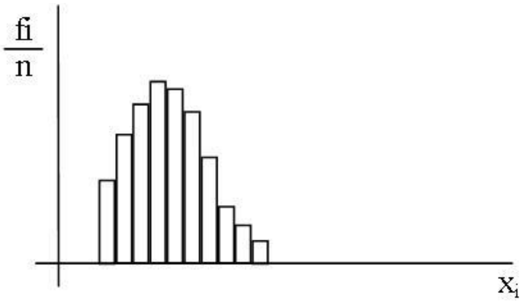
Se definió una variable continua como aquella que puede tomar un número infinito de valores de entre dos puntos fijos cualesquiera.

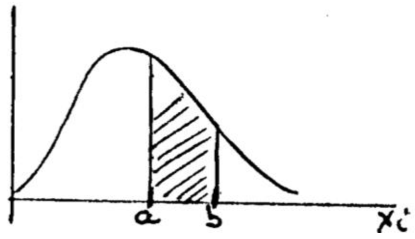
5.1 Función de densidad de probabilidad

Para ayudarnos a comprender la naturaleza de la distribución de una variable continua se debe pensar en el histograma de frecuencias relativas que origina una variable de este tipo e imagínese que el número de observaciones crece y que la amplitud de los intervalos de clase se hace muy pequeño. Es así como el polígono de frecuencia se transforma en una curva, que es la forma de representar las distribuciones de probabilidades continuas.

Esto se explica pensando que la superficie representa la suma de las probabilidades de todos los valores posibles de la variable aleatoria (condición de cierre)

Si se consideran dos puntos a y b sobre el eje x, el área bajo la curva limitada por a y b nos dará la P(a ≤ x ≤ b)





Para calcular probabilidades asociadas con sectores o intervalos es conveniente usar una expresión en x o una función no negativa de x, representada por $f(x)$ y que se conoce con el nombre de **función de densidad de probabilidad**.

Esta función densidad de probabilidad satisface:

a) $f(x) \geq 0$ para todo x

b) $\int_{-\infty}^{\infty} f(x)\,dx = 1$

c) Para cualquier a, b, tal que $-\infty < a < b < +\infty$ tenemos $P(a \leq X \leq b) = \int_a^b f(x)\,dx$

5.2 Distribución normal

La distribución de probabilidad conocida como distribución normal es, por la cantidad de fenómenos que explica, la más importante de las distribuciones estadísticas.

Para muchas clases de experiencias, en las ciencias biológicas, en las ciencias físicas, en la industria, etc., se encuentra que la distribución normal se aproxima a la distribución empírica obtenida en un gran número de repeticiones del experimento.

A la distribución normal también se la denomina con el nombre de campana de Gauss, pues al representar su función de probabilidad, ésta tiene forma de campana.

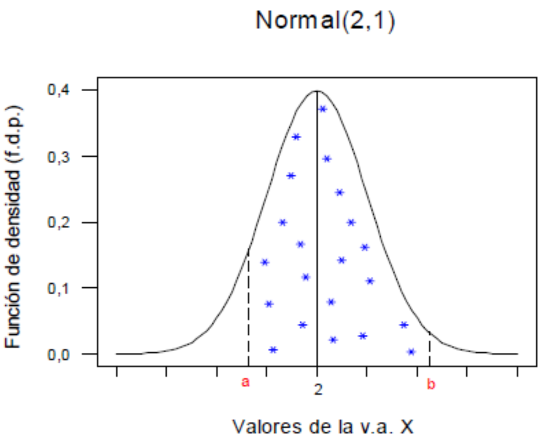
Gauss y Laplace, cuyos nombres lleva la ley, llegaron a ella estudiando la distribución de los errores de las observaciones. Bajo la influencia de sus trabajos se consideró mucho tiempo casi como un axioma que todas las distribuciones estadísticas se aproximarían a la normal si se dispusiera de un número suficientemente grande de observaciones.

Sin que hoy puedan aceptarse tales ideas, debe señalarse la extraordinaria frecuencia con que esta distribución se encuentra en las más variadas aplicaciones y también como otras pueden reducirse a ésta.

Para indicar que una variable aleatoria sigue una distribución normal de media $E(X)=\mu$ y desviación estándar $\sigma=\sqrt{V(X)}$ usaremos la expresión: *X ~ N (μ, σ)*

5.2.1 Definición de función de densidad

Diremos que la variable aleatoria X se distribuye normalmente con parámetros $E(X)=\mu$ y $\sigma=\sqrt{V(X)}$ (X ~ N (μ, σ)), si la función de densidad está dada por $f(x)=\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}$, para -∞< x < +∞ y la podemos representar para distintos valores de σ:

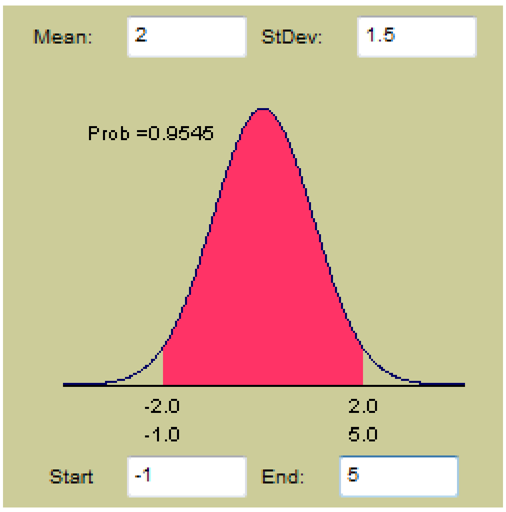


Cuanto más pequeña sea σ, tanto más concentrada resulta la masa de las distribución en el entorno del punto x = μ

Las características más generales de esta curva son:

- Tiene forma de campana, siendo simétrica con respecto a μ.
- La media, mediana y modo coinciden.
- Es asintótica con respecto al eje x.
- La distribución normal queda completamente determinada por los valores μ y σ es decir que existe una curva para cada par de estos valores.
- El rango intercuartil es aproximadamente igual a 1,33 σ

Con un valor de μ y de σ podemos determinar el área bajo la curva entre dos valores cualesquiera de la variable, (o sea, la probabilidad), para lo cual necesitamos hacer un cambio de escala. Este cambio de escala consiste en expresar la variable aleatoria X en términos de una variable Z, denominada variable aleatoria estandarizada.



Consultar el siguiente enlace:

http://psych.colorado.edu/~mcclella/java/normal/accurateNormal.html

En este enlace, a través de una sencilla aplicación, es posible comprobar a través de un applet, cómo la distribución normal representa un área bajo la curva.

Cambiando arriba los valores de la media y la desviación estándar, y abajo los valores entre los cuales queremos calcular la probabilidad, es decir, a qué porción de espacio bajo la curva normal corresponde la probabilidad buscada.

5.2.2 La distribución normal estándar

No existe una sola distribución de probabilidad normal, sino una "familia" de ellas. Cada una de las distribuciones puede tener una media (μ) o una desviación estándar distinta (σ). Por tanto, el número de distribuciones normales es ilimitado y sería imposible proporcionar una tabla de probabilidades para cada combinación de μ y σ.

Para resolver este problema, se utiliza un solo "miembro" de la familia de distribuciones normales, aquella cuya media es 0 y desviación estándar 1 que es la que se conoce como **distribución estándar normal**, de forma que todas las distribuciones normales pueden convertirse a la estándar, restando la media de cada observación y dividiendo por la desviación estándar. Primero, convertiremos la distribución real en una distribución normal estándar utilizando un valor llamado **estadístico Z** que será la distancia entre un valor seleccionado, designado X, y la media μ, dividida por la desviación estándar σ.

Entonces, dada una variable aleatoria continua X ~ N (μ, σ), la variable $Z=\dfrac{x-\mu}{\sigma}\approx N(0,1)$. Efectivamente:

$$E(Z)=E\left[\frac{x-\mu}{\sigma}\right]=\frac{1}{\sigma}E[x-\mu]=\frac{1}{\sigma}[E(X)-\mu]=0 \quad (\textbf{Esperanza matemática})$$

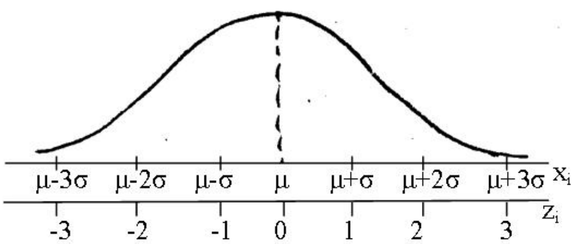
$$V(Z)=V\left[\frac{x-\mu}{\sigma}\right]=\frac{1}{\sigma^2}V[x-\mu]=\frac{1}{\sigma^2}V(X)=1 \quad (\textbf{Varianza})$$

La expresión de la **función de densidad** es la siguiente $f(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$

De esta manera, un valor Z mide la distancia entre un valor especificado de X y la media aritmética, en las unidades de la desviación estándar.

Por su condición de simetría, dos valores diferentes de x_i (o de z_i) que muestran la misma desviación en valor absoluto de μ tienen la misma densidad de probabilidad.

La transformación que sufre el eje de las x será:

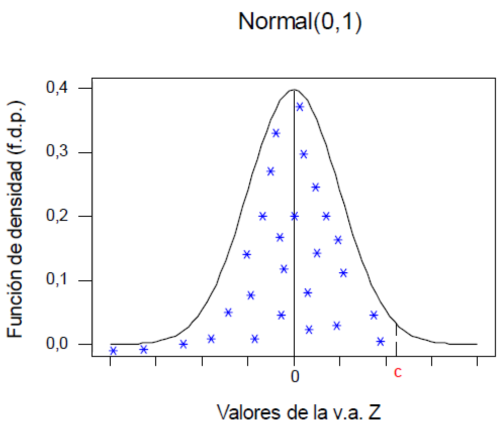


Al determinar el valor Z utilizando la expresión anterior, es posible encontrar el área de probabilidad bajo cualquier curva normal haciendo referencia a la distribución normal estándar en las tablas correspondientes.

Consultar: "Tabla de distribución normal estándar N (0, 1)" en:

http://www.vadenumeros.es/sociales/tabla-distribucion-normal-tipificada.htm

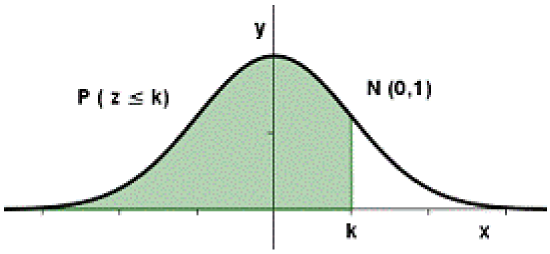
Dicha tabla proporciona la probabilidad de que la variable aleatoria normal estándar *Z* tome



un valor situado a la izquierda de un número c, P (Z<c).

En otras palabras, esta tabla nos da el valor del área encerrada por f(x) entre -∞ y c.

Si analizamos la tabla de Distribución Normal Estándar acumulada (- ∞ a *z*), en la parte superior se presenta el centésimo de *z* y a la izquierda el entero con el décimo correspondiente a cada valor de *z*.



z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549

Ejemplo 13:

Dada la distribución N (0,1) encontrar:

i) P (z ≤ 1)

ii) P (z ≥ 1)

iii) P (-2 ≤ z ≤ 0)

i) Buscando en Tabla se obtiene: P(z ≤ 1) = 0,8413

ii) Buscando en la Tabla se obtiene: P (z ≥ 1) = 1 – P (z ≤ 1) =
= 1- 0,8413 = 0,1587

iii) P (-2 ≤ z ≤ 0) = P (z ≤ 0) – P (z ≤ -2)

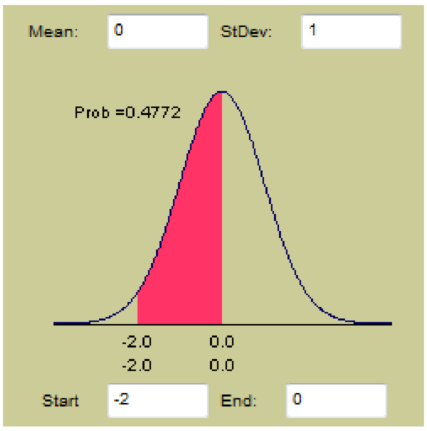
Buscando en tabla:

P (z≤0)=0,5

P (z≤-2)=1 - P(z≥2)=1 – 0,9772= 0,0228

Realizando:

P (-2 ≤ z ≤ 0) = P (z ≤ 0) – P (z ≤ -2)=
=0,5 – 0,0228= 0,4772



Consultar para otros ejemplos:

http://www.vadenumeros.es/sociales/manejo-tabla-normal.htm

Puede, además, interesar encontrar un valor en particular asociado con una probabilidad determinada.

Si analizamos la Tabla de Distribución Normal Acumulada de:

http://www.elosiodelosantos.com/sergiman/div/tablnorm.html

Si queremos encontrar z_0 siendo $P(z \leq z_0) = 0{,}1112 \Rightarrow z_0 = -1{,}22$

Para valores negativos de z

Normal	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5	0.496010644	0.492021686	0.488033527	0.484046563	0.480061194	0.476077817	0.47209683	0.468118628	0.464143607
-0.1	0.460172163	0.456204687	0.452241574	0.448283213	0.444329995	0.440382308	0.436440537	0.432505068	0.428576284	0.424654565
-0.2	0.420740291	0.416833837	0.412935577	0.409045885	0.405165128	0.401293674	0.397431887	0.393580127	0.389738752	0.385908119
-0.3	0.382088578	0.378280478	0.374484165	0.370699981	0.366928264	0.363169349	0.359423567	0.355691245	0.351972708	0.348268273
-0.4	0.344578258	0.340902974	0.337242727	0.333597821	0.329968554	0.32635522	0.32275811	0.319177509	0.315613697	0.312066949
-0.5	0.308537539	0.305025731	0.301531788	0.298055965	0.294598516	0.291159687	0.287739719	0.284338849	0.280957309	0.277595325
-0.6	0.274253118	0.270930904	0.267628893	0.264347292	0.2610863	0.257846111	0.254626915	0.251428895	0.24825223	0.245097094
-0.7	0.241963652	0.238852068	0.235762498	0.232695092	0.229649997	0.226627352	0.223627292	0.220649946	0.217695438	0.214763884
-0.8	0.211855399	0.208970088	0.206108054	0.203269392	0.200454193	0.197662543	0.194894521	0.192150202	0.189429655	0.186732943
-0.9	0.184060125	0.181411255	0.17878638	0.176185542	0.17360878	0.171056126	0.168527607	0.166023246	0.163543059	0.16108706
-1	0.158655254	0.156247645	0.15386423	0.151505003	0.14916995	0.146859056	0.1445723	0.142309654	0.14007109	0.137856572
-1.1	0.135666061	0.133499513	0.131356881	0.129238112	0.127143151	0.125071936	0.123024403	0.121000484	0.119000107	0.117023196
-1.2	0.11506967	0.113139446	0.111232437	0.109348552	0.107487697	0.105649774	0.103834681	0.102042315	0.100272568	0.098525329
-1.3	0.096800485	0.095097918	0.093417509	0.091759136	0.090122672	0.088507991	0.086914962	0.085343451	0.083793322	0.082264439
-1.4	0.080756659	0.079269841	0.077803841	0.07635851	0.0749337	0.07352926	0.072145037	0.070780877	0.069436623	0.068112118

6. Un poco de historia

Jacob Bernoulli descubrió que las frecuencias observadas se acercaban al verdadero valor previo de su probabilidad al hacer crecer el número de repeticiones del experimento. Pero él quería encontrar una prueba científica que no sólo probara que al aumentar el número de observaciones de la muestra se podía estimar la probabilidad

auténtica con el grado de precisión deseado en cada ocasión, sino que permitiera calcular explícitamente cuántas observaciones eran necesarias para garantizar esa precisión de que el resultado queda dentro de un intervalo predeterminado alrededor de la verdadera solución.

El experimento que consiste repetir una prueba con la misma probabilidad de éxito un número grande de veces recibió el nombre de "experimento de Bernoulli" y, más adelante, tras la creación del concepto de variable aleatoria, la variable que contabiliza el número de éxitos en *N* pruebas se llamó 'Bernoulli' o 'binomial'.

Bernoulli era consciente de que, en situaciones reales y cotidianas, la certeza absoluta, es decir, la probabilidad 1, es imposible de alcanzar. Por eso introdujo la idea de la "certeza moral": para que un resultado fuese moralmente cierto, debía tener una probabilidad no menor que 0.999, mientras que un resultado con probabilidad no mayor que 0.001 se consideraría "moralmente imposible". Fue para determinar la certeza moral de un suceso para lo que Bernoulli formuló su teorema, la ley de los Grandes Números.

Para tener una idea intuitiva de este concepto, Bernoulli propuso el siguiente ejemplo: una urna con 30.000 bolas blancas y 20.000 negras, aunque el observador no lo sabe, pues lo que quiere es determinar la proporción entre bolas blancas y negras, sacando una de cada vez, anotando el resultado (éxito si es blanca y fracaso si es negra) y volviéndola a introducir en la urna.

Sea N el número de observaciones, X el número de éxitos y $p = \frac{r}{r+s}$ la probabilidad de éxito en cada prueba, siendo r el número de bolas blancas y s el de bolas negras. El teorema de Bernoulli afirma que dada cualquier pequeña fracción ε (que su descubridor siempre escribía en la forma $\frac{1}{r+s}$ y dado cualquier número entero positivo grande c, se puede hallar un número $N = N(c)$ tal que la probabilidad de que $\frac{X}{N}$ difiera de p no más de ε es mayor que c veces la probabilidad de que $\frac{X}{N}$ difiera de p más de ε. Veamos todo esto con notación matemática:

$$P\left\{\left|\frac{X}{N} - p\right| \leq \varepsilon\right\} > c \,.\, P\left\{\left|\frac{X}{N} - p\right| > \varepsilon\right\}$$

O en notación más actual:

$$\forall \varepsilon > 0 \; \forall c \in Z^{+} \; \exists N \; tl \; que \; P\left\{\left|\frac{X}{N} - p\right| > \varepsilon\right\} < \frac{1}{c+1}$$

Bernoulli tomó como ejemplo para aclarar este problema c=1.000 y obtuvo como resultado que eran necesarias 25.550 observaciones. La intuición le decía que no hacían falta tantas y, por ello, lo intentó con otros valores de c. Desilusionado por sentir que había fracasado en su intento de cuantificar la certeza moral, Bernoulli no incluyó en su libro las aplicaciones prometidas.

El que sí lo hizo fue su sobrino Niklaus Bernoulli (1687–1759), que aplicó el resultado de su tío a registros de 14.000 nacimientos y llegó a la inesperada conclusión de que la frecuencia de nacimientos de niños es mayor que la de niñas, en la proporción de 18:17. Los resultados tanto de Bernoulli como de su sobrino fueron confirmados años después por Laplace.

De esta manera, gracias a Bernoulli, se introdujo en la teoría de la probabilidad la ley de los Grandes Números, uno de los conceptos más importantes en cálculo de probabilidades, muestreos, etc… y con amplias aplicaciones en muchos campos de la estadística y de las matemáticas y la ciencia en general. Esta ley además será objeto de conversaciones entre matemáticos en los siglos venideros, estando sujeta a constantes estudios, mejoras y ampliaciones hasta prácticamente nuestros días.

A raíz del descubrimiento de la Ley de los Grandes Números, se planteó el problema de estimar la suma de un subconjunto de elementos de una expresión de carácter binomial. La dificultad era calcular una probabilidad que Bernoulli ya había dejado indicada y demostrada y era la probabilidad de que el número de éxitos de un suceso de probabilidad p y n pruebas estuviera entre A y B.

Según Bernoulli esta probabilidad era $\sum_{A<k<B}^{1} \binom{m}{k} p^{k} \, (1-p)^{m-k}$ y la dificultad que entrañaba esta operación era el cálculo del número combinatorio $\binom{m}{k} = \frac{m!}{k!(m-k)!}$

Jacob consiguió realizar ciertas estimaciones de dudosa precisión, pero que al no manejar números muy grandes fueron suficientes para demostrar sus teoremas.

Posteriormente, De Moivre intentó ser algo más preciso y demostró que para B constante $m! \approx B \,.\, e^{-m} \,.\, m^{m+\frac{1}{2}}$ Con el fin de determinar el valor de esa constante, construyó la siguiente expresión logarítmica $\ln B = 1 - \frac{1}{12} + \frac{1}{360} - \frac{1}{1260} + \frac{1}{1680} - \ldots$ y concluyó según sus cálculos que B era igual a 2,5074.

Este resultado parecía convincente, pero De Moivre no quedó satisfecho al no poder vincularlo con ninguna constante matemática conocida. Por ese motivo pidió consejo y ayuda a su buen amigo **James Stirling** (1692-1770) quien demostró que $B = \sqrt{2\pi}$

Tras la obtención de este importante dato, De Moivre calculó una tabla m! y obtuvo un resultado que dice lo siguiente:

$$P\left\{X = \frac{n}{2} + t\right\} \approx P\left\{X = \frac{n}{2}\right\}.e^{-(2t^2/n)} = \frac{2}{\sqrt{2\pi n}}\,.\,e^{-(2t^2/n)}$$

De Moivre dibujó la gráfica de esta curva, introduciendo el importantísimo concepto de “distribución normal” y demostró que esta curva es simétrica en torno al máximo y que los puntos de inflexión están a distancia $\frac{1}{2}.\sqrt{n}$ de este máximo. Por tanto, aquí está el Teorema Central del Límite que se conoce como el hecho de que el promedio de una muestra aleatoria independiente convenientemente normalizada sea aproximadamente normal.

Ahora De Moivre ya estaba en condiciones de resolver el problema planteado por Bernoulli en la Ley de los Grandes Números:

$$\sum P\left\{X = \frac{n}{2} + t\right\} = \frac{2}{\sqrt{2\pi n}} \int_0^k .e^{-(2t^2/n)} dt$$

Para finalizar, De Moivre encontró que $\sqrt{n}$ era la unidad cuya distancia del centro debe ser medida.

Así, la precisión de una estimación de la probabilidad aumenta igual que la raíz cuadrada del número de experimentos; en otras palabras, De Moivre acababa de descubrir la importancia de la varianza.

De Moivre repitió el experimento de Bernoulli y obtuvo que basten 6.498 observaciones. Aunque mejoró el método, no llegó a reconocer la importancia del descubrimiento de la curva de la normal y no pudo aplicar su resultado a otros problemas.