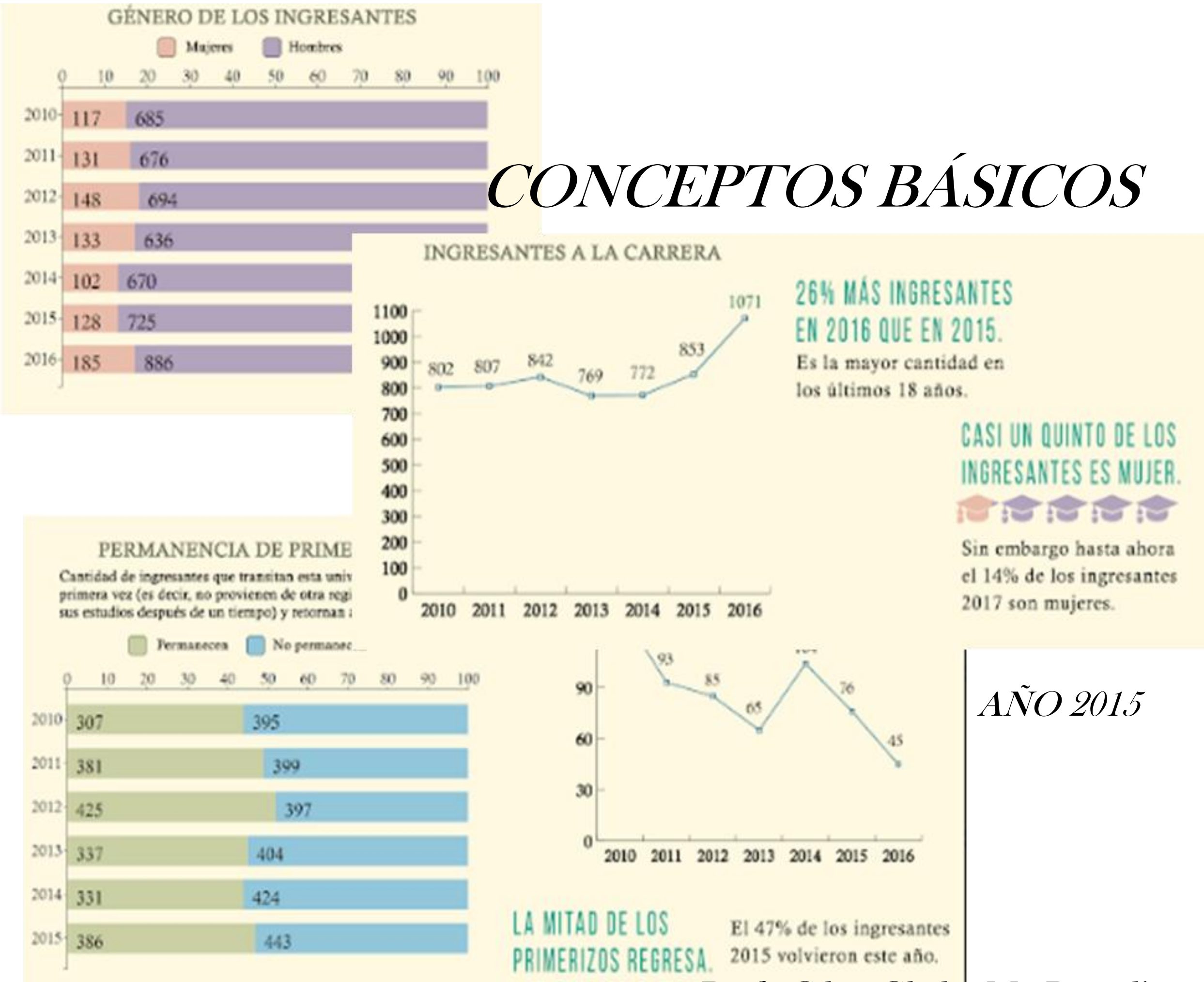
Universidad Tecnológica Nacional
Facultad Regional Córdoba
Ingeniería En Sistemas De Información

PROBABILIDADES Y ESTADÍSTICAS

CONCEPTOS BÁSICOS



AÑO 2015

Prof. Cdra. Gladys M. Rouadi

La Estadística constituye una disciplina científica que trata de la selección, análisis y uso de datos con el fin de resolver problemas. A toda persona, tanto en su ejercicio profesional como en su actividad diaria en contacto con diferentes medios, se le ofrece información en forma de datos. Consecuentemente, algunos conocimientos de Estadística le serán de utilidad a la población en general, pero en particular en conocimiento estadístico será de vital importancia para ingenieros de todas las especialidades, científicos y administradores, debido a que manejan y analizan datos cotidianamente. En consecuencia las herramientas básicas de la Estadística les resultan de gran importancia a la hora del ejercicio profesional.
Las aplicaciones de la Probabilidad y la Estadística son numerosas en todos los casos de la ciencia aplicada en donde existan variaciones y donde las conclusiones acerca de un sistema están basadas en datos observados.
Por Estadística y Probabilidad se entiende los métodos para describir y modelar la variabilidad, además de permitir la toma de decisiones cuando la variabilidad está presente.
Del diseño a la producción, los procesos tienen que ser permanentemente mejorados. Con sus conocimientos técnicos y dotados de habilidades estadísticas básicas para la recolección y representación gráfica de datos, ingenieros y científicos podrán desenvolverse eficientemente. *Agradezco a los integrantes de la Cátedra y a los alumnos que colaboraron en la detección de errores que permiten año tras año mejorar al presente material.*

Probabilidades y Estadística: conceptos básicos. 1º ed.-Córdoba. ROUADI, Gladys Margarita. Eudecor. 2013. ISBN 978-987-1536-38-2. Fecha de catalogación: 12/04/2013

CAPITULO Nº 4: Teoría de Probabilidades

Objetivos Específicos
Que el estudiante:

Conozca *los conceptos más relevantes de la teoría de probabilidades.*
Reconozca *la necesidad de su estudio como instrumento para medir la incertidumbre.*
Comprenda *su aplicación en la construcción de modelos que describan la realidad y posibiliten su análisis.*
Reconozca y comprenda *sus aportes en la inferencia estadística.*

Contenidos

1. *Conceptos generales.*
2. *Incertidumbre y experimento aleatorio.*
3. *Espacios probabilísticos.*
4. *Eventos.*
  4.1. *Concepto.*
  4.2. *Eventos mutuamente excluyentes.*
  4.3. *Eventos no mutuamente excluyentes.*
  4.4. *Eventos colectivamente exhaustivos.*
  4.5. *Eventos no colectivamente exhaustivos.*
5. *Interpretación de la probabilidad de un hecho.*
6. *Teorías probabilísticas.*
  6.1. *Historia de la teoría de probabilidades.*
  6.2. *Principio de la razón insuficiente.*
  6.3. *Teoría frecuencial.*
  6.4. *Teoría subjetivista.*
  6.5. *Axiomatización de la probabilidad.*
    6.5.1. *Axiomas y propiedades para la familia de eventos.*
    6.5.2. *Axiomas y propiedades para la probabilidad de los eventos.*
7. *Probabilidad total. Regla aditiva especial.*
8. *Probabilidad condicional.*
9. *Probabilidad compuesta o conjunta. Regla multiplicatoria general.*
10. *Probabilidad marginal o individual.*
11. *Dependencia e independencia estadística.*
12. *Teorema o regla de Bayes.*
13. *Aplicaciones de la teoría de probabilidad: procesos estocásticos.*

1. *Conceptos Generales*

Hasta aquí hemos estudiado la Estadística Descriptiva (Recolección, Organización, Presentación y Descripción De Datos De Población o De Muestra). Pero los datos de muestra, generalmente no son estudiados por ellos mismos, sino para revisar un conocimiento anterior de quien toma decisiones sobre los estados de la naturaleza o para generalizar acerca de un parámetro de población desconocido.

Así, el problema que vincula los resultados muestrales con los valores poblacionales, se denomina Inferencia Estadística y comprende fundamentalmente dos tipos de problemas: *La Estimación Y La Docimasia De Hipótesis.*

Los datos de muestra pueden ayudar a reducir nuestra incertidumbre sobre los estados de la naturaleza, pero rara vez pueden eliminar la incertidumbre en una situación dada de toma de decisiones. Así, la toma de decisiones con o sin datos de muestra se asocia siempre con un grado de incertidumbre. *La Estadística, Como Un Método De Toma De Decisiones Frente A La Incertidumbre, Se Basa En La Teoría De Probabilidades, Porque La Probabilidad Es A La Vez El Lenguaje y La Medida De La Incertidumbre y Los Riesgos Asociados Con Ella.*

Entonces, con la base de datos de una muestra, se aprenderá la forma de obtener conclusiones acerca de la población de la cual se obtuvo la muestra. La Probabilidad es fundamental para la *Inducción* estadística. Probabilidad y Estadística guardan íntima relación.

En la Estadística, se introduce la mano en una caja opaca de bolillas blancas y de color, se examina el puñado de bolillas extraídas y se trata de contestar la siguiente pregunta: *¿Qué hay en la caja?*

En la Probabilidad, se mira a una caja transparente, se cuentan las bolillas blancas y de color en ella, se mezclan cabalmente, y después a ciegas se saca un puñado, sin abrir los ojos, se predice cuántas bolillas de cada clase se tienen en la mano.

Se aprenderá a contestar la pregunta acerca de la caja negra únicamente después de que se haya estudiado la probabilidad, pues se utilizará este estudio para elegir la muestra, para hacer el cálculo acerca de la población (el contenido de la caja opaca) y para decidir qué grado de confianza puede depositarse en la estimación.

La Teoría de Probabilidad, además de ser un elemento imprescindible para la *Estimación Estadística* y la Toma de Decisiones a través de la *Docimasia de Hipótesis*, es también de gran importancia en todos aquellos problemas donde interviene la *Incertidumbre.*

2. *Incertidumbre y Experimento Aleatorio*

Un experimento aleatorio o estocástico es aquel que conduce a dos ó más resultados posibles. Uno de estos posibles resultados se presentará, pero no podemos asegurar cuál, entonces, estos resultados son inciertos.

Así, echar una moneda o un dado, son experimentos aleatorios, porque en cada caso el proceso produce más de un resultado posible.

Un experimento puede consistir de sólo una prueba, o estar conformado de un conjunto de pruebas realizadas bajo las mismas condiciones.

3. Espacios Probabilísticos

El Espacio Probabilístico, también llamado Espacio Muestral o Espacio de Muestra, es el conjunto de todos los resultados posibles de un experimento aleatorio. Es el equivalente al Conjunto Universal, visto en la Teoría de Conjuntos, es decir a aquel conjunto formado por todos los elementos que intervienen en el tema de interés. Se simboliza por Ω (omega). Los resultados posibles de un espacio probabilístico o de muestra se llaman elementos o puntos de muestra.

Cualquier prueba de un experimento produce un resultado que corresponde exactamente a un elemento de dicho espacio.

Ejemplo:.

1- Tirada de un dado: $\Omega = (1, 2, 3, 4, 5, 6)$

2- Tirada de una moneda: $\Omega = (c, s)$

3- Tirada conjunta de un dado y una moneda

$\Omega = ?$

$\Omega_1 = (c, s)$ $\Omega_2 = (1, 2, 3, 4, 5, 6)$

Entonces, Ω puede considerarse como el conjunto de productos cartesianos, es decir el conjunto cuyos elementos son todos los pares ordenados, cuyo primer componente pertenece a Ω_1, y cuyo segundo componente pertenece a Ω_2.

$$\Omega_1 \times \Omega_2 = (c, s) \times (1, 2, 3, 4, 5, 6)$$

$\Omega = \{ (c,1); (c,2); (c,3); (c,4); (c,5); (c,6); (s,1); (s,2); (s,3); (s,4); (s,5); (s,6) \}$

Un espacio probabilístico o muestral con un pequeño número de elementos o puntos de muestra puede representarse bien gráficamente por un Sistema de Coordenadas Cartesianas Ortogonales o por un Diagrama de Árbol.

Sistema de Coordenadas Cartesianas

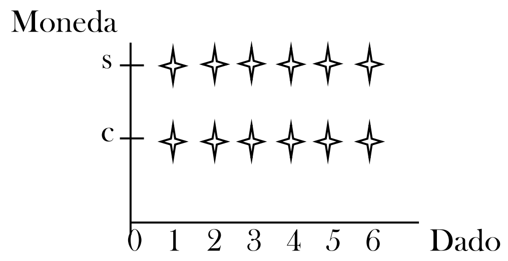
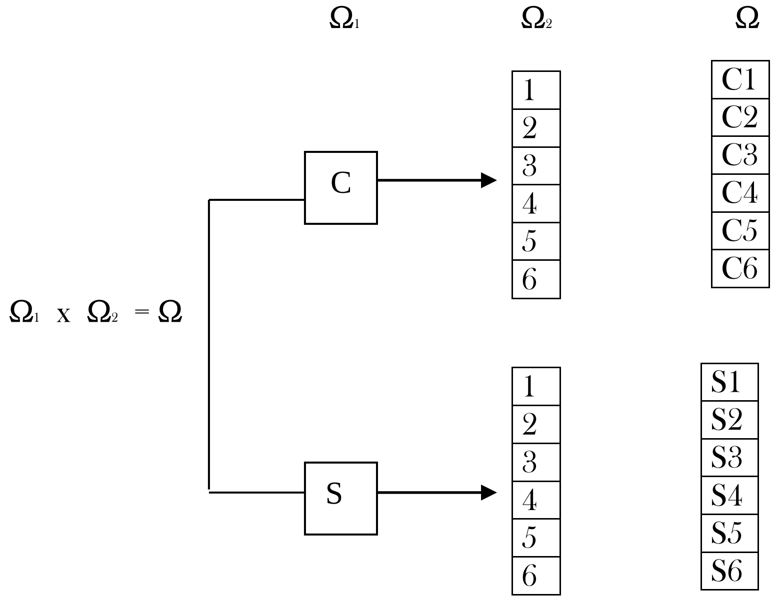


Diagrama de Árbol



Vemos que los diagramas de árbol proporcionan una forma organizada de enumerar todos los elementos posibles de un espacio de muestra de modo que no falte ninguno.

Cualquiera sea la forma en que se escriba o represente diagramáticamente, los resultados de Ω, deben ser *Mutuamente Excluyentes y Colectivamente Exhaustivos.*

Por ejemplo, en la tirada de una moneda, tenemos, según vimos, dos resultados posibles Ω = (c,s); pero cada lanzamiento de la moneda, producirá, ya sea cara o sello, pero no ambos simultáneamente. Es decir, que cada punto del espacio probabilístico correspondiente a este experimento, es mutuamente excluyente en una prueba.

En todos los casos, el espacio probabilístico está constituido por un número finito de puntos, no obstante, pueden considerarse también:

- Infinidad numerable de puntos. Por ejemplo, lanzar en forma repetida un dado hasta obtener el número 4. Los resultados posibles, podrían representarse por 1, 2, 3,..., n,..., según que se obtenga el número 4 en el primer lanzamiento, en el segundo, en tercero,..., en el enésimo,..., etc., respectivamente.
- Infinidad no numerable de puntos. Por ejemplo, tirar a un blanco que tiene formas de círculo, es decir una infinidad no numerable de puntos.

4. Eventos

Un evento, también llamado hecho, es un subconjunto de un espacio probabilístico o de muestra. Los diferentes eventos o hechos serán representamos por las letras mayúsculas, tales como E_1, E_2 ,...

1- Un evento o hecho E definido sobre un espacio de muestra Ω, se dice que *es un evento simple, elemental, o fundamental*, si contiene exactamente un punto de muestra en Ω.

Así, en la tirada de un dado, Ω = (1, 2, 3, 4, 5, 6), entonces cada uno de los elementos de Ω es un evento simple. Es decir que habrá tantos eventos elementales, simples o fundamentales como resultados posibles tenga el experimento aleatorio.

A partir de esto, podemos definir al Espacio Probabilístico como el conjunto de eventos simples, elementales, o fundamentales.

2- Un evento o hecho E definido sobre un espacio de muestra Ω, se llama *un compuesto o evento (hecho) compuesto o simplemente evento* (hecho), si contiene más de un punto de muestra un Ω.

Así, cuando se echa un dado y definimos E_1 = (1, 3, 5); E_2 = (2, 4, 6); entre otros, son compuestos.

Los eventos incluyen particularmente los siguientes subconjuntos:

1- *Un Evento Elemental*, es también un evento, en el sentido de que es un subconjunto que contiene un solo punto del espacio probabilístico.

2- *El Espacio Probabilístico* Ω, es un evento, en el sentido de que es un subconjunto que contiene todos los eventos elementales de Ω. Es el *Evento Cierto.*

3- Un subconjunto que no contiene puntos de Ω, *es el Evento Imposible*, dado que no puede ocurrir nunca.

Ejemplo en la tirada de un dado donde Ω = (1, 2, 3, 4, 5, 6)

1- Evento imposible: salida del número 7.

$$E_0 = \varnothing$$

2- Eventos simples

$$E_1 = (1)$$
$$E_2 = (2)$$
$$\vdots$$
$$E_6 = (6)$$

3- Eventos compuestos

E_7 = (1, 2, 3)

E_8 = (2, 5, 6)

E_9 = (5, 6)

4- Evento cierto

Ω = (1, 2, 3, 4, 5, 6)

Es importante establecer aquí, dos aspectos en relación a los eventos:

1- Un evento ha ocurrido, si al menos uno de sus elementos se ha presentado.

Por ejemplo, en la tirada de un dado, definimos E_1 = Salida de número par, es decir E_1 = (2, 4, 6). Si tiramos el dado, y sale el número 6, el evento E_1 ha ocurrido.

2- Los eventos y sólo los eventos poseen probabilidad asociada.

Si ahora consideramos a dos eventos definidos sobre el mismo espacio probabilístico, pudiendo extender las definiciones a más de dos eventos, podremos clasificarlos de la siguiente manera:

4.1. Eventos Mutuamente Excluyentes

Dos eventos o hechos definidos sobre un mismo espacio probabilístico o de muestra, son Mutuamente Excluyentes, Mutuamente Exclusivos, Disjuntos ó Desunidos, si no hay puntos de muestras en común, es decir, si su intersección es igual al vacío.

Simbólicamente:

Si $E_1 \cap E_2 = \varnothing$ Entonces los eventos son Mutuamente Exclusivos.

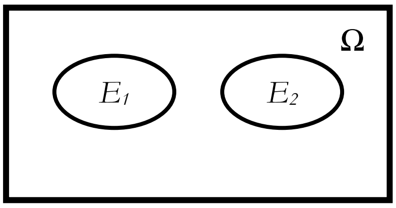
Ejemplo:

En el experimento tirada de un dado, el espacio probabilístico queda definido como:

$$\Omega = (1, 2, 3, 4, 5, 6)$$

Si definimos los eventos:
E_1 = (1, 3) E_2 = (2, 4), entonces $E_1 \cap E_2 = \varnothing$

Gráficamente, usamos un Diagrama de Venn:



4.2. Eventos No Mutuamente Excluyentes

Dos eventos o hechos definidos sobre un mismo espacio probabilístico o de muestra, son No Mutuamente Excluyentes, No Mutuamente Exclusivos, No Disjuntos ó No Desunidos, si hay puntos de muestras en común, es decir, si su intersección es distinta al vacío.

Simbólicamente:

Si $E_1 \cap E_2 \neq \emptyset$ Entonces los eventos son no Mutuamente Exclusivos.

<u>Ejemplo:</u>

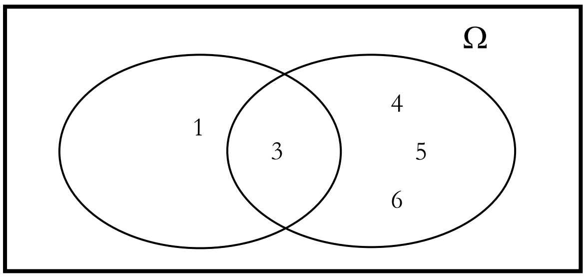
En el experimento tirada de un dado, el espacio probabilístico queda definido como:

$$\Omega = (1, 2, 3, 4, 5, 6)$$

Si definimos los eventos:

$E_1 = (1,3)$ $E_2 = (3, 4, 5, 6)$, entonces $E_1 \cap E_2 \neq \emptyset$, ya que $E_1 \cap E_2 = (3)$

Gráficamente, usamos un Diagrama de Venn:



Los eventos no Mutuamente Exclusivos pueden ser:

Dependientes: Cuando la ocurrencia o no ocurrencia de un evento en cualquier prueba, afecta la probabilidad de otros eventos en las pruebas siguientes.

Independientes: Cuando la ocurrencia o no ocurrencia de un evento en cualquier prueba, **no** afecta la probabilidad de otros eventos en las pruebas siguientes.

4.3. Eventos Colectivamente Exhaustivos

Se dice que dos hechos ó eventos, definidos sobre el mismo espacio de muestra, son *Colectivamente Exhaustivos*, si su unión es igual al espacio de muestra.

Simbólicamente:

Si $E_1 \cup E_2 = \Omega$ Entonces los eventos son Colectivamente Exhaustivos.

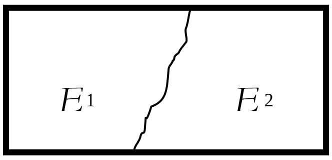
<u>Ejemplo:</u>

$\Omega = (1, 2, 3, 4, 5, 6)$
$E_1 = (1, 3, 5)$
$E_2 = (2, 4, 6)$

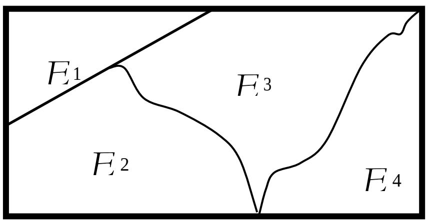
E_1 y E_2 son Colectivamente Exhaustivos, porque $E_1 \cup E_2 = \Omega$.

Además, en este caso, E_1 *y* E_2 son Mutuamente Exclusivos, porque $E_1 \cap E_2 = \varnothing$*,* lo que hace que el conjunto (E_1 , E_2) sea una partición de Ω.

Gráficamente:



Entonces, diremos que, si n hechos forman una partición de Ω, los n hechos son Mutuamente Exclusivos y Colectivamente Exhaustivos.



Obsérvese:

- Si dos ó más eventos o hechos son Mutuamente Exclusivos, no pueden ocurrir juntos, por lo que, como máximo, ocurrirá uno.

- Si dos ó más eventos o hechos son Colectivamente Exhaustivos, por lo menos ocurrirá uno de ellos.

- Así, si un conjunto de eventos o hechos son mutuamente exclusivos y colectivamente exhaustivos, ocurrirá exactamente uno de los hechos.

4.4. Eventos No Colectivamente Exhaustivos

Se dice que dos hechos ó eventos, definidos sobre el mismo espacio de muestra, son *No Colectivamente Exhaustivos*, si su unión *NO* da por resultado al espacio de muestra.

Simbólicamente:

Si $E_1 \cup E_2 \neq \Omega$ Entonces los eventos son No Colectivamente Exhaustivos.

Ejemplo:

$\Omega = (1, 2, 3, 4, 5, 6)$
$E_1 = (1, 3)$
$E_2 = (2, 4)$

E_1 y E_2 son No Colectivamente Exhaustivos, porque $E_1 \cup E_2 \neq \Omega$.

5. Interpretación de la Probabilidad de un hecho

A partir de lo visto en los apartados anteriores, concluimos que la probabilidad de un evento o hecho dado es una expresión de su posibilidad de ocurrencia.

Además, una probabilidad es un número que varía de cero, para el evento o hecho que no puede ocurrir, es decir el Evento Imposible, a uno, para el evento o hecho que es cierto que ocurra, es decir, el Evento Cierto (Recordemos que ***Un evento ha ocurrido, si al menos uno de sus elementos se ha presentado***).

Es decir que tenemos dos certezas, que el evento imposible no ocurrirá, por lo tanto su probabilidad de ocurrencia es 0, y que el evento cierto siempre ocurrirá, por lo tanto su probabilidad de ocurrencia es 1.

Ejemplo:

Dado el experimento Tirada de un dado, donde $\Omega = (1, 2, 3, 4, 5, 6)$

Si definimos:

E_1 = (Salida del nº 7) = Evento imposible, dado que el N° 7 no es un resultado posible,
por lo tanto $P(E_1) = 0$

E_2 = (1, 2, 3, 4, 5, 6)= Evento cierto, y con seguridad ocurrirá uno de los resultados posibles, por lo tanto $P(E_2) = 1$

Pero, ¿Cómo asignar probabilidades para aquellos eventos que no sean el imposible o el cierto?

La respuesta no es única, pues depende de la situación en la que nos encontremos. Es por ello, que existen tres escuelas de pensamiento principales para darnos una solución a nuestra necesidad, llamadas Teorías Probabilísticas.

6. Teorías Probabilísticas

6.1. Historia de la Teoría de Probabilidades

La probabilidad matemática tiene sus orígenes en los juegos de azar, principalmente con dados y cartas. Los primeros estudios "científicos" sobre fenómenos aleatorios se centraban en dos problemas:

1. Contabilizar el número de posibles resultados al lanzar un dado varias veces.
2. Distribuir las ganancias entre jugadores cuando el juego se interrumpía antes de finalizar, conocido como el "problema de las apuestas"

En estos problemas trabajaron Richard de Fournival (1200-1250), que determinó 216 combinaciones posibles para el primer caso y calcula los diferentes valores para la suma de los tres dados. Por otro lado, Luca Pacioli (1445-1517), se encarga del segundo caso, dando una solución que con el tiempo se determinó como incorrecta.

La primera obra fue el Libro de los Juegos de Azar, de Girolamo Cardano (1501-1576), escrito en 1565 y publicado en 1663. El autor, jugador empedernido, escribe más bien un tratado de juegos de azar, dedicando una mínima parte al estudio del azar. Trabajó con los conceptos de la teoría clásica de probabilidad pero no los definió.

También Galileo Galilei (1564-1642) abordó problemas sobre dados. Su obra *Sobre la Puntuación en Tiradas de Dados,* calculaba el número de resultados posibles tirando tres dados, a partir de $6^3 = 216$. Sin embargo, su principal contribución fue la creación de la teoría de la medida de los errores. Consideró que los errores de medida eran inevitables y los clasificó en *errores sistemáticos*, debidos a los métodos y herramientas de medida; y los *errores aleatorios*, que varían impredeciblemente de una medida a otra. Esta concepción sigue aún vigente. Al analizar las propiedades de los errores estableció que son más frecuentes los errores pequeños que los grandes; que los errores por defecto son tan frecuentes como los errores por exceso; y que la mayoría de las mediciones se agrupan alrededor del verdadero valor. Con esto contribuyó al desarrollo de la teoría de probabilidades, pero también sentó bases para el nacimiento de la Estadística.

El desarrollo de la teoría de la probabilidad experimentó un gran avance en Francia a mediados del siglo XVII con la correspondencia entre Blaise Pascal (1623-1629) y Pierre de Fermat (1601-1665) durante 1654. Antoine Gombaud, caballero de Méré, filósofo y literato, además de jugador compulsivo, le pide a Pascal que resuelva el problema del reparto de apuestas. Pascal y Fermat lo resolvieron correctamente por medios diferentes pero equivalentes. El acierto de ambos provenía de considerar que el reparto de las apuestas debe hacerse en función de la probabilidad de ganar que tuviese cada jugador en el momento de interrumpirse el juego. Lo que Pascal no pudo es extender el procedimiento al caso en que hubiera tres o más jugadores.

Once años más tarde, en 1665, Pascal en su *Tratado sobre el Triángulo Aritmético,* realiza la contribución más importante en el campo de la combinación.

Pascal aplicó los razonamientos probabilísticos sobre la toma de decisiones a la Teología y trató de demostrar la existencia de Dios. Su argumento fue el siguiente: Dios existe o no existe: si no existe, da igual creer en Él que no creer; si existe, creer que no existe provoca la condenación eterna, mientras que creer trae la salvación. Como la salvación es preferible a la condenación (en términos probabilísticos, la ganancia es mayor), una persona "razonable" actuará como si Dios existiera, aunque crea que la probabilidad de que exista es pequeña.

Los trabajos de Pascal y Fermat fueron continuados por el científico holandés Christiaan Huygens (1629-1695). Su interés por la probabilidad nació en 1655 durante el transcurso de un viaje a París, donde coincidió con otros científicos y discutió con ellos el problema del reparto de apuestas. En 1656 publica el tratado *Sobre los Cálculos en los Juegos de Azar.* Huygens introducía el concepto de esperanza matemática para variables aleatorias que toman dos o tres valores, definida como la ganancia media si se repitiera el juego muchas veces. Aportó su solución al problema del reparto de apuestas, muy similar a Pascal, pero fue capaz de extenderlo al caso de tres jugadores.

Hacia el último cuarto del siglo XVII, existía una importante magnitud de conocimientos sobre sucesos aleatorios, con un importante planteo de problemas planteados con sus resoluciones.

Las primeras investigaciones de la probabilidad no centradas en juegos de azar, se realizaron en Demografía, a través de los estudios de un comerciante inglés John Graunt (1620-1675), quien en 1662 se propuso encontrar un método preciso para estimar la edad media de los habitantes de Londres a través de la edad de defunción, introduciendo el concepto de "frecuencia de un suceso". También demostró que en Londres la proporción de nacimientos de niños y niñas no era igual y elaboró la primera tabla de mortalidad. Comprendió que cuántas más observaciones hacía más precisos eran los resultados, anticipando el principio estadístico de estabilidad de las medias.

Las ideas de Graunt fueron tomados por William Petty (1623-1687), quien elaboró un análisis comparativo entre la mortalidad en Londres y en París, basándose en los datos de los hospitales de caridad, y por el astrónomo Edmund Halley (1656-1742), quien presentó en 1693 una tabla de mortalidad de la ciudad de Breslau (Alemania). Este último prefería ciudades pequeñas con pocos movimientos migratorios. Introdujo además, el concepto de longitud de vida tal que la frecuencia con que se superaba o no se alcanzaba era la misma, es decir la mediana. En sus trabajos puede encontrarse también las bases de los teoremas de la suma y la multiplicación de probabilidades y de la ley de los Grandes Números. Su obra tuvo gran influencia en Demografía y Seguros

6.2. Teoría Clásica o Principio de la Razón Insuficiente

El primero en dar la definición clásica de probabilidad fue Jakob Bernoulli (1654-1705) en su obra *El Arte de Predecir* , publicada en 1713, muy influenciada por los trabajos de Graunt y Petty, que habían demostrado las ventajas de incluir en sus tablas no sólo los números absolutos, sino también las proporciones respecto del total. Más adelante, el matemático francés exiliado en Inglaterra Abraham De Moivre (1667-1754) aceptó la definición de Bernoulli y la reformuló en términos modernos: *"una fracción en la que el numerador es igual al número de apariciones del suceso y el denominador es igual al número total de casos en los que ese suceso pueda o no pueda ocurrir. Tal fracción expresa la probabilidad de que ocurra el suceso".* La definición clásica de probabilidad, en su forma actual, está basada en el concepto de equiprobabilidad, es decir, resultados igualmente probables de un experimento aleatorio.

En este supuesto, no existe elemento alguno, o razón alguna, para preferir uno cualquiera de los resultados posibles sobre cualquier otro, entonces todos deben ser tratados como si tuvieran la misma oportunidad de ser elegidos. Es decir, todos los resultados posibles deben tener los mismos pesos o probabilidades de ocurrencia.

Así, si un dado perfecto es echado, debe considerarse que hay igual probabilidad que salga cualquiera de los seis números, como consecuencia, la probabilidad de que salga cualquier número, por ejemplo un tres, será 1/6.

Las anteriores observaciones nos conducen a la sencilla y directa interpretación clásica de la probabilidad.

Si el espacio de muestra de un experimento tiene $N(\Omega)$ resultados igualmente probables, y si un evento o hecho, definido en este espacio de muestra, tiene $n(E)$ elementos, la probabilidad queda definida:

$$P(E) = \frac{n(E)}{N(\Omega)} = \frac{\text{Número de casos favorables al evento } E}{\text{Número de casos posibles en } \Omega}$$

Por ejemplo, si se echa un dado y se define:

A = (Salida de un número impar) = (1,3,5)

Entonces, $P(A) = \frac{3}{6} = \frac{1}{2}$

B = (Que el número sea mayor a 4) $= (x > 4) = (x \geq 5) = (5,6)$

Entonces, $P(B) = \frac{2}{6} = \frac{1}{3}$

C = (Salida de un 1 ó 2) = (1,2)

Entonces, $P(C) = \frac{2}{6} = \frac{1}{3}$

Debemos señalar que la teoría clásica, en el supuesto de resultados igualmente probables, depende del razonamiento lógico.

Así, generalmente no encontramos ninguna dificultad si tenemos una moneda bien equilibrada, un dado no cargado, una ruleta honesta, o cualquier otro experimento cuyos resultados son igualmente probables y pueden ser deducidos por lógica. Sin embargo, hay razón para creer que el mundo real no tiene cosas tales. Por lo tanto, el supuesto de perfección causará asignaciones de probabilidad ligeramente incorrectas.

Y, ¿Qué podemos decir de una moneda desequilibrada?, ¿de un dado cargado?, ¿de una ruleta arreglada?

En cada uno de estos casos, el enfoque clásico de asignar probabilidades iguales daría cálculos de probabilidad incorrectos.

Por fortuna, cuando fracasa un razonamiento a priori, debido a la falta de probabilidades iguales, podemos consultar la Teoría de la Probabilidad por Frecuencia Relativa.

6.3. Teoría Frecuencial o Teoría de la Probabilidad por Frecuencia Relativa

Basándose en los trabajos de Graunt y Petty, Bernoulli introdujo también el concepto de probabilidad frecuentista o estadística: asignar como probabilidad de un suceso el resultado que se obtendría si el proceso se repitiera en condiciones similares un número grande de veces, ideando la Ley de los Grandes Números, en base a que era consciente de que las frecuencias observadas se acercaban a un cálculo previo de su probabilidad al aumentar el número de repeticiones del experimento. .

Entonces, los teóricos de frecuencia relativa afirman que el único procedimiento válido para determinar probabilidades de hechos, es por experimentos repetitivos.

Por ejemplo, cuando se echa una moneda, ¿Cuál es la probabilidad de que salga anverso? El teórico de frecuencia relativa abordaría el problema echando realmente la moneda, por ejemplo 100 veces, en las mismas condiciones, y calculando después la proporción de veces que sale anverso.

Supongamos que la moneda saca anverso 45 veces de 100 tiradas, entonces la razón 45/100 se usa como una estimación de la probabilidad de anverso, *P(A)* de esta moneda. Un momento de reflexión mostrará que aún si la moneda estuviera perfectamente equilibrada, la estimación se acercaría a la razón verdadera (Probabilidad) de 1/2 cuando el número de pruebas aumenta.

Esta exposición, nos conduce a la siguiente interpretación de probabilidad en términos de frecuencia relativa: si un experimento es ejecutado n veces en las mismas condiciones y hay n_i resultados, $n_i \leq n$, en que ocurrió un hecho, entonces una estimación de la probabilidad de ese hecho es la razón n_i / n.

Además, la estimación de la probabilidad de un hecho n_i / n se acerca a un límite, la verdadera probabilidad del hecho, cuando n aumenta sin límite, es decir,

$$P(E) = Lím_{n \to \infty} n_i / n$$

Naturalmente, en la práctica nunca podemos obtener la probabilidad de un hecho dado por este límite; sólo podemos buscar una estimación próxima de $P(E)$ basada en n grande.

Por comodidad, trataremos la estimación de $P(E)$ como si fuera realmente $P(E)$, escribiendo la definición de probabilidad por frecuencia relativa:

$$P(E) = n_i / n = h_i \qquad \text{y} \qquad \begin{array}{l} 0 \leq n_i / n \leq 1 \\ 0 \leq P(E) \leq 1 \end{array}$$

Definiendo $P(E)$ como un límite cuando n se aproxima a infinito se destaca que la probabilidad supone el concepto a largo plazo.

La Teoría Clásica y la Teoría Frecuencial se llaman definiciones objetivas de probabilidad.

La definición Clásica es objetiva en el sentido de que se basa en la deducción de un conjunto de supuestos. La definición de Frecuencia Relativa es objetiva porque la probabilidad de un hecho es determinada por repetidas observaciones empíricas.

6.4. Teoría Subjetiva o Teoría Personalista

Un objetivista se encuentra muy a gusto hablando de probabilidad en relación con el lanzamiento de una moneda o la fabricación de un producto en masa. Sin embargo, podría sentirse impotente respecto al problema de hechos únicos, hechos que ocurren una sola vez o que no pueden ser sometidos a experimentos repetitivos. Como resultado, una clase grande de problemas se encuentra más allá del alcance del objetivista.

Esta limitación de la Teoría de la Frecuencia Relativa y el supuesto clásico de probabilidades iguales, ha favorecido el punto de vista personalista sobre la probabilidad.

Así, en el segundo cuarto del siglo XX surgió una nueva interpretación, llamada subjetiva, según la cual la probabilidad mide el grado de creencia de un individuo en la verdad de una proposición, variando entre 0, el individuo cree que es falso, a 1, cree que es cierto.

Fue propuesta por primera vez por el filósofo Frank P. Ramsey en su libro *Los Fundamentos de las Matemáticas* de 1931, y el primer matemático que la adoptó fue el estadístico italiano Bruno de Finetti en 1937.

El teórico personalista o subjetivo, considera la probabilidad como una medida de confianza personal en una proposición particular, es decir asigna un peso entre 0 y 1 a un evento o hecho, según su grado de creencia en su posible ocurrencia.

Por ejemplo, si tiene doble confianza en la ocurrencia del hecho A que el hecho B, y si A y B son los únicos hechos posibles, asigna los valores:

$$P(A) = 2/3 \qquad \text{y} \qquad P(B) = 1/3.$$

La interpretación subjetiva de la probabilidad es más amplia que la Frecuencial, pues mientras que ésta sólo funciona con experimentos que se puedan repetir un número grande de veces, aquélla se puede aplicar a cualquier tipo de proposiciones. Además, mientras que los frecuentistas consideran que la probabilidad de un suceso o hecho es siempre una constante (que se aproxima repitiendo el proceso), para los subjetivistas la probabilidad de un suceso puede, y debe, variar en función de la nueva información recibida respecto del suceso, manera de proceder que se ajusta más al método científico.

Una crítica que se ha hecho es la supuesta arbitrariedad con que se asignan probabilidades a sucesos o hechos. La defensa aduce que las normas de la moral y la coherencia obligan a quien asigna la probabilidad a actuar del modo más objetivo posible y no de forma caprichosa. Además, se han hecho esfuerzos por convertir esta noción intuitiva en demostraciones formales.

Es muy utilizada en los diseños de modelos probabilísticos de la física cuántica, y sus técnicas se han aplicado con éxito recientemente para filtrar el spam del correo electrónico legítimo.

6.5. La Axiomatización de la Probabilidad

Como era necesaria una formalización más grande de los conceptos probabilísticos y hubo que esperar el desarrollo de la Teoría de Conjuntos y de la Medida que tuvo lugar a finales del siglo XIX y comienzos del XX, debidos principalmente a Georg Cantor (1845-1918), Émile Borel y Henri Lebesgue (1875-1941). Ya en 1909, Borel consideró la importancia de la teoría general de la medida para la construcción de la fundamentación de la teoría de la probabilidad, pero no fue hasta 1933 cuando N. Kolmogorov se propuso construir una teoría de la probabilidad de una manera, basándose en axiomas fundamentales.

La construcción axiomática de la teoría de la probabilidad procede de las propiedades fundamentales de la probabilidad observadas en los ejemplos que ilustran las definiciones Clásica y Frecuencial. Así, la definición axiomática las incluye como casos particulares y supera las carencias de ambas.

De esta manera, la probabilidad pudo desarrollarse como una teoría completamente lógica al mismo tiempo que siguió permitiendo resolver los problemas aplicados de las ciencias modernas y la tecnología.

Para ello se han creado Axiomas y Propiedades tanto para la Familia de Eventos como para las Probabilidades, que nos permitirán su cálculo cuando ejecutamos operaciones con eventos, tales como la unión, la intersección, la complementación, la diferencia.

6.5.1. Axiomas y Propiedades Para La Familia De Eventos (F)

Axioma F_1**:** Si $\varnothing$ es el evento imposible y Ω es el evento cierto, ambos pertenecen a la familia de eventos (F).

Simbólicamente

$$\varnothing \text{ y } \Omega \in F \,.$$

Axioma F_2: Dado un conjunto numerable de eventos A_1, A_2, A_3,..., la intersección de ese conjunto numerable es un evento.

Simbólicamente

$$\text{Si } A_1, A_2, A_3, \ldots\ldots \in F \Rightarrow \bigcap_{i=1}^{\infty} A_i \in F$$

En particular, si a partir de un Ω dado, definimos dos eventos A y B, entonces el resultado de $A \cap B$ también es un Evento.

Axioma F_3**:** Dado un conjunto numerable de eventos A_1, A_2, A_3,..., la unión de ese conjunto numerable es un evento.

Simbólicamente

$$\text{Si } A_1, A_2, A_3, \ldots\ldots \in F \Rightarrow \bigcup_{i=1}^{\infty} A_i \in F$$

En particular, si a partir de un Ω dado, definimos dos eventos A y B, entonces el resultado de $A \cup B$ también es un Evento.

Axioma F_4**:** Si A es un evento y $\bar{A}$ es su complemento, entonces $\bar{A}$ es un evento.

Simbólicamente

$$\text{Si } A \in F \Rightarrow \overline{A} \in F$$

Propiedad : Si A y B son eventos, entonces la diferencia entre A y B también es un evento.

Simbólicamente

$$\text{Si } A \text{ y } B \in F \Rightarrow A - B \in F$$

6.5.2. Axiomas y Propiedades Para La Probabilidad De Los Eventos

Axioma P_1: Si A es un evento tiene asociada una probabilidad

Simbólicamente

$$\text{Si } \; A \in F \Rightarrow \exists \text{ para } A, \text{ un número } P(A)$$

Axioma P_2: Positividad ó Ley de No Negatividad. La Probabilidad de un hecho en un espacio de muestra es no negativa.

Simbólicamente

$$P(A) \geq 0 \quad \forall \text{ evento } A$$

Axioma P_3: Certidumbre. La probabilidad de todo espacio de muestra es 1.

Según vimos, un evento ha ocurrido, si al menos uno de sus elementos se ha presentado. Cuando se efectúa un experimento, estamos seguros de que por lo menos uno de los resultados, es decir uno de los puntos de muestra de Ω debe ocurrir, y, por lo tanto, la probabilidad asociada con Ω es 1.

Axioma P_4: Ley de Probabilidad Total ó Regla Aditiva Especial

Si $A_1, A_2, A_3, \ldots$, es un conjunto numerable de eventos, finito ó infinito, mutuamente excluyente, entonces la probabilidad de la unión de dichos conjuntos es igual a la suma de sus probabilidades.

Simbólicamente

$$\text{Si} \quad A_1, A_2, A_3, \ldots\ldots \in F \;\; y \; A_i \cap A_j = \varnothing \qquad i \neq j \quad i, j = 1,2,3,\ldots$$

$$P\left(\bigcup_{I=1}^{\infty}\right) = \sum_{1=1}^{\infty} P\left(A_I\right) = P(A_1) + P(A_2) + P(A_3) + \ldots$$

A partir de estos cuatro Axiomas, es posible establecer un conjunto de propiedades básicas:

Propiedad P_5:

$$Si\ A\ y\ B \in F \;\; / \; A \subset B \Rightarrow P(B - A) = P(B) - P(A)$$

<u>Ejemplo:</u>

Si en la tirada de un dado, definimos los siguientes eventos:
A = (2,3)
B = (2,3,4,5)

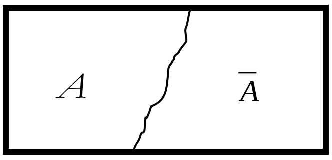
Entonces: *P(A)* = 2/6 y *P (B)* = 4/6

Luego: $B - A = (4{,}5)$ y $P(B - A) = \frac{4}{6} - \frac{2}{6} = \frac{2}{6}$

Propiedad P_6: Ley de complementación.

$$\text{Si } A \in F \Rightarrow P(\overline{A}) = 1 - P(A)$$

Sea $\overline{A}$ el complemento de A en Ω. Aquí $\overline{A}$ y A son hechos complementarios, porque como lo indica la figura son mutuamente excluyentes y colectivamente exhaustivos.



Se deduce que $P(A \cup \overline{A}) = P(A) + P(\overline{A}) = 1$, de donde: $P(\overline{A}) = 1 - P(A)$

Ejemplo:

Supongamos que se echan dos dados al mismo tiempo, ¿Cuál es la probabilidad de obtener un doble?

Para resolver este problema, observemos que los hechos "no obtener un doble" y "obtener un doble", son complementarios. Además, el espacio de muestra de este experimento, consta de 36 resultados posibles:

	1	2	3	4	5	6
1	***1,1***	1,2	1,3	1,4	1,5	1,6
2	2,1	***2,2***	2,3	2,4	2,5	2,6
3	3,1	3,2	***3,3***	3,4	3,5	3,6
4	4,1	4,2	4,3	***4,4***	4,5	4,6
5	5,1	5,2	5,3	5,4	***5,5***	5,6
6	6,1	6,2	6,3	6,4	6,5	***6,6***

Entre ellos, hay seis puntos que corresponden al hecho "obtener un doble" (diagonal principal).

Designemos a este conjunto por A, entonces: $P(A) = \frac{6}{36} = \frac{1}{6}$

La probabilidad buscada, de "no obtener un doble", es por lo tanto:

$$P(\overline{A}) = 1 - \frac{6}{36} = \frac{30}{36} = \frac{5}{6}$$

Propiedad P_7: Si Ω es cualquier espacio de muestra y P es cualquier función de probabilidad definida en Ω, entonces $P(\varnothing) = 0$.

Este teorema afirma que si el conjunto de hechos es el conjunto vacío, el hecho es una imposibilidad para la cual la probabilidad de ocurrencia es 0. El conjunto vacío no contiene punto de muestra, por lo tanto, no puede asignársele ningún peso. Por ejemplo, la probabilidad de obtener un 7 al echar un dado ordinario es 0. No hay tal número en el dado.

Propiedad P_8: La probabilidad de un evento es número que va de 0 a 1.

$$\text{Si } A \in F, \ entonces \quad 0 \leq P(A) \leq 1$$

Propiedad P_9: Dados dos eventos *A* y *B*, de tal manera que *A* está contenido en *B*, es decir que *A* es un subconjunto propio de *B* (todos los elementos de *A* están en *B*, pero existen elementos de *B* que no están en *A*), entonces la probabilidad de *A* es menor que la probabilidad de *B*.

$$\text{Si } A \ y \ B \in F \ / \ A \subset B \Rightarrow P(A) \langle \ P(B)$$

Ejemplo:

A = (2, 3)	P (*A*) = 2/6 = 0,333		
B = (2, 3, 4, 5)	P (*B*) = 4/6 = 0,666	→	P(*A*) < P(*B*)

Propiedad P_{10}: Si A_1, A_2, A_3,..., es un subconjunto finito numerable de eventos, no necesariamente mutuamente excluyentes,

$$P(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i)$$

Esta propiedad se denomina ***Desigualdad de Boole***, y puede considerarse como una generalización de la probabilidad total.

7. Probabilidad Total. Regla Aditiva Especial

La Ley de Probabilidad Total, referida a la probabilidad de Unión de Eventos, fue vista en el Axioma P_4 y la Propiedad P_{10}, de los Axiomas y Propiedades para la Probabilidad de los Eventos.

Desarrollaremos con más detalle a esta Ley.

Recordemos el axioma P_4 : Si $A_1, A_2, A_3, \ldots\ldots.$; es un conjunto numerable de eventos (finito o infinito), *Mutuamente Excluyente*, entonces la probabilidad de la unión de dichos conjuntos $(\bigcup_{i=1}^{\infty} A_i)$ es igual a la suma de sus probabilidades.

$$\text{Si } A_1, A_2, A_3; \ldots \in \text{F}$$

$$\text{y } A_i \cap A_i = \varnothing \qquad i \neq j \qquad i, j = 1;\ 2;\ 3;\ldots\ldots$$

Entonces

$$P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) = P(A_1) + P(A_2) + P(A_3) + ...$$

Si en particular consideramos que $A \; y \; B$, definidos en Ω, son hechos *Mutuamente Excluyentes* o *Disjuntos*, es decir que se verifica que: $A \cap B = \varnothing$, entonces, según esta ley la probabilidad de la unión de ambos es igual a la suma de sus probabilidades. En Símbolos:

$$P(A \cup B) = P(A) + P(B).$$

Por otro lado, la Propiedad P_{10}, rezaba: "Si $A_1, A_2, A_3, ...$, es un subconjunto infinito numerable de eventos, *NO NECESARIAMENTE Mutuamente Excluyentes*, entonces la probabilidad de la unión de dichos conjuntos es *menor o igual* a la suma de sus probabilidades".

Esta propiedad nos plantea la solución para el cálculo de la Probabilidad de la Unión de eventos o hechos, tanto en el caso en que los mismos sean Mutuamente Excluyentes, o No Mutuamente Excluyentes.

Según lo indicado por la Ley de Probabilidad Total, vista en el apartado anterior, la probabilidad de la unión de eventos Mutuamente Excluyentes es igual a la suma de sus probabilidades.

Por lo tanto, la Probabilidad de la unión de eventos ***No Mutuamente Excluyentes***, será menor a la suma de sus probabilidades.

En símbolos:

Si $A_1, A_2, A_3, ... \in F \;\; y \; A_i \cap A_j \neq \varnothing$, entonces:

$$P\left(\bigcup_{I=1}^{\infty}\right) \langle \sum_{1=1}^{\infty} P(A_I) \langle \; P(A_1) + P(A_2) + P(A_3) + ...$$

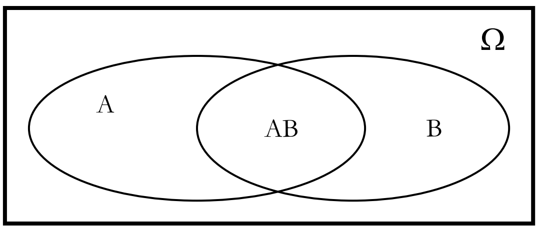
En particular, dados dos eventos A y B, definidos en Ω, *No Mutuamente Excluyentes*, la probabilidad de la unión de dichos eventos es igual a la suma de sus probabilidades menos la probabilidad de la intersección de dichos eventos.

En símbolos:

Si $A \; y \; B \in F \;\; y \;\; A \cap B \neq \varnothing$, entonces:

$$P(AUB) = P(A) + P(B) - P(AB)$$

Gráficamente:



Esto ocurre, por el siguiente motivo:

Sabemos que la *P (AUB)* es la suma de las probabilidades de todos los puntos individuales de *AUB.*

Ahora bien, *P(A) + P (B)* es el total de la suma de las probabilidades de los puntos de *A* y la suma de las probabilidades de los puntos de *B.* Sin embargo, $A \cap B \neq \varnothing$, entonces *P(A) + P(B)* incluye las probabilidades de todos los puntos de la intersección $A \cap B$ dos veces.

Obviamente, si deducimos de esta suma la $P(A \cap B)$ una vez, el resultado es el total de probabilidades de todos los puntos de *AUB*, cada uno de los cuales se toma una sola vez.

A este resultado se lo conoce también como ***Regla Aditiva General***, pues es una generalización de la Probabilidad Total, ya que si los eventos son mutuamente excluyentes la probabilidad de la intersección será 0, pues $A \cap B = \varnothing$, mientras que si los eventos son no mutuamente excluyentes, tendrá un valor distinto (>) a 0, pues $A \cap B \neq \varnothing$.

<u>Ejemplo:</u>

Si se extrae una carta de un mazo de cartas francesas y se pide la probabilidad de:

a) Que aparezca un as o una reina
b) Que aparezca un corazón o un diamante
c) Que aparezca una carta superior al valor 10 o un tres
d) Que aparezca una REINA o un CORAZÓN

Recordemos que un mazo de cartas francesas tiene 52 cartas y consideremos además, que todas son igualmente posibles, de manera que el espacio probabilístico está compuesto por 52 eventos elementales, cada uno de ellos con una probabilidad asociada de 1/52.

Existen cuatro figuras distintas que son: Corazón, Diamante, Trébol y Pique, cada una con 13 cartas. Las cartas con valores superiores a diez son: el as, el rey, la reina y la sota.

Tendremos entonces:

a) A = (Que aparezca un AS) $\Rightarrow P(A) = 4/52$

B = (Que aparezca una REINA) $\Rightarrow P(B) = 4/52$ y dado que $A \cap B = \varnothing$

$$P(A \cup B) = P(A) + P(B) = \frac{4}{52} + \frac{4}{52} = \frac{8}{52}$$

b) C = (Aparición de un CORAZÓN) $\Rightarrow P(C) = 13/52$

D = (Aparición de un DIAMANTE) $\Rightarrow P(D) = 13/52$ y dado que $A \cap B = \varnothing$

$$P(C \cup D) = P(C) + P(D) = \frac{13}{52} + \frac{13}{52} = \frac{26}{52}$$

c) E = (Aparición de una carta > al valor 10) =(As, Rey, Reina, Sota) $\Rightarrow P(E) = 16/52$

F = (Aparición de un Tres) $\Rightarrow P(F) = 4/52$ y dado que $E \cap F = \varnothing$

$$P(E \cup F) = P(E) + P(F) = \frac{16}{52} + \frac{4}{52} = \frac{20}{52}$$

d) G = {Que aparezca una REINA} $\Rightarrow P(G) = \dfrac{4}{52}$

H = {que aparezca un CORAZÓN} $\Rightarrow P(H) = \dfrac{13}{52}$

$G \cap H$ = {REINA y CORAZÓN} $\Rightarrow P(G \cap H) = \dfrac{1}{52}$

Y $P(G \cup H) = P(G) + P(H) - P(G \cap H) = \dfrac{4}{52} + \dfrac{13}{52} - \dfrac{1}{52} = \dfrac{16}{52}$

Ejemplo:

Una caja contiene 200 piezas idénticas de máquina, de las cuales 100 son producidas por la máquina A, 60 por la máquina B y 40 por la máquina C. Si una pieza es escogida al azar de la caja,

6.6. ¿Cuál es la probabilidad de que haya sido producida por la máquina A ó B?

$$P(A \cup B) = P(A) + P(B) = \frac{100}{200} + \frac{60}{200} = 0{,}80$$

6.7. ¿Cuál es la probabilidad de que haya sido producida por la máquina B ó C?

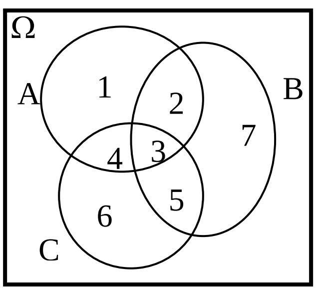
$$P(B \cup C) = P(B) + P(C) = \frac{60}{200} + \frac{40}{200} = 0{,}50$$

A los fines de una mayor claridad plantearemos el tema como unión de más de dos conjuntos. Así definimos los siguientes conjuntos:

A = (1, 2, 3, 4)

B = (2, 3, 5, 7)

C = (3, 4, 5, 6)



Entonces:

$A \cup B \cup C$ = (1, 2, 3, 4, 5, 6, 7), que obtuvimos de la siguiente forma:

$A \cup B \cup C = (A) + (B) + (C) - (A \cap B) - (A \cap C) - (B \cap C) + (A \cap B \cap C)$

$A \cup B \cup C$ =(1, 2, 3, 4) + (2, 3, 5, 7) + (3, 4, 5, 6) - (2, 3) - (3, 4) - (3, 5) + (3)

$A \cup B \cup C$ = (1, 2, 3, 4, 5, 6, 7)

Ejemplo:

Hay 300 estudiantes en un Centro de Estudios que ofrece entre otros, tres lenguas orientales: Chino (*A*), japonés (*B*), birmano (*C*). El director del Centro ha proporcionado los siguientes datos sobre la matrícula de alumnos para estas lenguas:

Lengua	N° de Estudiantes
A	100
B	80
C	50
$A \cap B$	16
$A \cap C$	10
$B \cap C$	8
$A \cap B \cap C$	1

Supongamos que un estudiante es escogido al azar de este Centro, ¿Cuál es la probabilidad de que estudie por lo menos una de las tres lenguas?

$$P(A \cup B \cup C) = \frac{100}{300} + \frac{80}{300} + \frac{50}{300} - \frac{16}{300} - \frac{10}{300} - \frac{8}{300} + \frac{1}{300} = \frac{197}{300} = 0.66$$

(197: Número total de estudiantes matriculados en por lo menos una clase de idiomas).

Hasta aquí hemos aprendido la solución para encontrar probabilidades cuando unimos eventos.

Para descubrir la forma para el cálculo de probabilidades en la intersección, debemos previamente analizar otros conceptos tales como Probabilidad Condicional

8. Probabilidad Condicional

En algunas circunstancias especiales es posible disponer de información que reduce el espacio probabilístico original a un subconjunto; es decir, se podrá trabajar con la probabilidad de una parte más bien que con todo el espacio original.

Evidentemente, la probabilidad de un evento acerca del cual contamos con alguna información, será diferente a la situación para la cual no tengamos ninguna información.

Por ejemplo, un estudiante seleccionado entre aquellos que han obtenido notas altas en Matemáticas, tendrá mayores posibilidades de obtener igual resultado en un curso de Estadística, en relación a un estudiante seleccionado de entre la totalidad de los que concurren a dicho curso.

En este ejemplo, la atención se centra en la probabilidad de un evento establecido a partir de un subconjunto pudiendo ser mayor, igual, o menor que en el espacio original.

Cada uno de los subconjuntos, es un espacio probabilístico reducido, y está especificado por nuevas condiciones que aparecen planteadas por la información adicional que se tiene en relación al espacio total.

Las probabilidades asociadas a los eventos definidos en las subpoblaciones se denominan *Probabilidades Condicionales*.

Para explicar esta sección es necesario recordar la definición Clásica De Probabilidad. Supongamos, que el espacio probabilístico Ω, que corresponde al conjunto original cuenta con *N* eventos elementales, todos con la misma probabilidad.

Sean:

A, un evento que tiene "a" eventos elementales favorables

B, otro evento que tiene "b" eventos elementales favorables y

$A \cap B$, con "c" eventos elementales favorables.

Ha ocurrido el evento *B*, y se pide la probabilidad de que también haya ocurrido el evento *A*, que en símbolos se expresa *P(A/B)* y se lee: "Probabilidad de *A* dado *B*", ó bien "Probabilidad de *A* con la condición de que haya ocurrido *B*".
El nuevo espacio probabilístico será ahora *B*, con "b" eventos elementales de los cuales "c" serán favorables al evento $A \cap B$. Es decir, en este nuevo espacio probabilístico, será:

$$P(A/B) = \frac{c}{b} = \frac{Eventos\ \ Favorables\ \ a\ \ A \cap B}{Eventos\ \ Favorables\ \ a\ \ B\,(espacio\ muestral\ reducido)}$$

Con respecto al espacio probabilístico general, Ω, esta probabilidad es:

$$P(A/B) = \frac{c}{b} = \frac{c/N}{b/N} = \frac{P(A \cap B)}{P(B)}$$

<u>Ejemplo:</u>

Supongamos que los 300 empleados de una empresa son clasificados transversalmente como personal gerencial o no gerencial, y como graduados universitarios o no universitarios, como sigue:

	Universitarios (*B*)	No Universitarios (*D*)	Total
Gerencial (*A*)	25	5	30

No Gerencial (*C*)	75	195	270
Total	100	200	300

TABLA I

Si se hace una elección al azar de los empleados, ¿cuál es la probabilidad de que sea un empleado gerencial?

A = (Empleado gerencial) y $P(A) = \dfrac{a}{N} = \dfrac{30}{300} = 0{,}10$

Y un graduado universitario?

B = (Graduado Universitario) y $P(B) = \dfrac{b}{N} = \dfrac{100}{300} = 0{,}33$

Suponiendo ahora que de la elección resulta un graduado universitario, ¿cuál es la probabilidad de *A* en vista de esta información adicional?

Esta información adicional, en efecto reduce el espacio de muestra Ω, a la subpoblación de graduados universitarios *B,* con b = 100 eventos elementales favorables.

Si nos interesamos por la probabilidad de *A* dado *B*, debemos conocer los elementos que pertenecen a $A \cap B$, que para el ejemplo es c = 25, entonces:

$$P(A/B) = \frac{c}{b} = \frac{25}{100} = 0{,}25$$

El resultado anterior puede ser comprobado considerando que P(*A/B*) es igual a la razón de la proporción de todos los empleados que son personal gerencial y que son graduados universitarios, P(*A*∩*B*), a la proporción de empleados que son graduados universitarios, P(*B*) con referencia al espacio muestral original.

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{c/N}{b/N} = \frac{25/300}{100/300} = 0{,}25 \text{, como antes.}$$

Podemos generalizar ahora la *Ley de Probabilidad Condicional:* "La probabilidad de que ocurra un hecho *A,* dado que otro hecho *B* ha ocurrido, se llama probabilidad condicional de *A dado B* y se expresa como sigue:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad ; \quad P(B) \rangle 0$$

Similarmente, la probabilidad condicional de B dado A, se define como:

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \quad ; \quad P(A) \rangle 0$$

Por ejemplo, en el caso anterior, ¿Cuál es la probabilidad de observar un empleado universitario, dado que la elección al azar produce un empleado gerencial?

$$P(B/A) = \frac{25/300}{30/300} = 0{,}83$$

Este resultado indica que, en general, P (A/B) ≠ P (B/A), porque aunque

P ($A \cap B$) = P ($B \cap A$), todas las veces,

P(A) puede diferir a menudo de P (B),

O, expresado de otra manera: n ($A \cap B$) = n ($B \cap A$) = c, todas las veces, pero n(A) = a, puede diferir a menudo de n (B) = b.

Esta observación significa que podemos interesarnos por diferentes espacios de muestra reducidos para el mismo conjunto de intersecciones.

Observaciones:

1) Todos los eventos están asociados con algún espacio probabilístico. Así, P(A) = P(A/Ω), pero en general omitimos Ω, ya que tácitamente lo damos por sobreentendido. El símbolo de condicionalidad se emplea solamente cuando existe algún subconjunto de Ω, que contenga todos los resultados de un cierto experimento. En tal caso tenemos P(A/A) = 1.

2) Las probabilidades indicadas en el cociente de las fórmulas, [P (A) y P (B)], corresponden a eventos planteados en el espacio original. Se puede obtener el mismo resultado si primero obtenemos el espacio probabilístico reducido, ya que las probabilidades asignadas a los eventos definidos en dicho espacio, son proporcionales a las asignadas a los eventos en el espacio probabilístico original Ω. Es decir, la suma de las probabilidades de los eventos elementales en el espacio probabilístico reducido será igual a la unidad.

3) La probabilidad del evento en el espacio reducido debe ser mayor que 0, ya que de otro modo la probabilidad condicional no estaría definida. En otras palabras, siempre debe ser distinta de 0 la probabilidad del denominador de las fórmulas, es decir P(A) y P (B). La P(A/B) puede ser igual ó diferente de P(A), según veremos más adelante.

9. Probabilidad Compuesta o Conjunta. Regla Multiplicatoria General

Para desarrollar el presente tema, continuaremos con el ejemplo dado en el apartado anterior, donde hemos transformado la información en términos de Probabilidad.

	Universitarios (B)	No Universitarios (D)	Total
Gerencial (A)	0,083	0,017	**0,10**

No Gerencial (*C*)	0,25	0,65	**0,90**
Total	**0,333**	**0,667**	**1**

Los valores del cuerpo de la tabla se llaman *Probabilidades Conjuntas*, porque cada una de ellas es la probabilidad de la ocurrencia conjunta, o simultánea, de dos hechos.

Así,

$P(A \cap B) = 0{,}083$, es la probabilidad de ocurrencia conjunta de empleado Gerencial y Universitario.

$P(D \cap C) = 0{,}65$, es la probabilidad de ocurrencia conjunta de No Universitario y No Gerencial.

Y así sucesivamente.

"La ley que rige la Probabilidad de ocurrencia conjunta de *A* y *B* se da por:

$$P(A \cap B) = P(B)\,P(A/B) \quad (1)$$

$$P(B \cap A) = P(A)\,P(B/A) \quad (2)$$

A esto lo hemos obtenido despejando de:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad ; \quad Para\ (1)$$

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \quad ; \quad Para\ (2)$$

En el ejemplo dado:

$$P(A \cap B) = P(A/B)\,P(B) = 0{,}25 \text{ x } 0{,}33 = 0{,}083$$

$$P(B \cap A) = P(B/A)\,P(A) = 0{,}83 \text{ x } 0{,}10 = 0{,}083$$

Observemos algunas inferencias de la Regla Multiplicatoria General:

1- El orden carece de importancia en el conjunto de intersecciones, porque

$$A \cap B = B \cap A$$

Esta propiedad de intersección, da el siguiente resultado:

$$P(A \cap B) = P(B \cap A)$$

2- La Regla multiplicatoria general puede ser generalizada para más de dos hechos.

$$P(ABC) = P(A)\,P(B/A)\,P(C/AB)$$

3- $P(A/B)$ raramente es igual que $P(A \cap B)$, puede ser enteramente diferente. Además, $P(A/B) = P(A)$ sólo en una condición especial, que se observará después.

<u>NOTA</u>: $P(A \cap B)$ puede también escribirse como $P(AB)$.

10. <u>Probabilidad Marginal o Individual</u>

Las *Probabilidades Marginales ó Individuales*, se obtienen sumando cada fila ó cada columna, y reciben este nombre simplemente porque se encuentran en los márgenes del cuadro.

	Universitarios (*B*)	No Universitarios (*D*)	Total
Gerencial (*A*)	0,083	0,017	**0,10**
No Gerencial (*C*)	0,25	0,65	**0,90**
Total	**0,333**	**0,667**	**1**

TABLA II

Obsérvese que una Probabilidad Marginal es la suma de un conjunto de Probabilidades Conjuntas.

Así:

$P(A) = P(AB) + P(AD) = 0{,}083 + 0{,}017 = 0{,}10$; es la probabilidad de un empleado Gerencial.

$P(B) = P(BA) + P(BC) = 0{,}083 + 0{,}25 = 0{,}333$; es la probabilidad de un empleado Universitario.

Y así sucesivamente.

Desarrollados y ejemplificados los conceptos de Probabilidad Condicional, Conjunta y Marginal, obtendremos de la TABLA I:

1- Probabilidades Condicionales

a- $P(A/B) = \dfrac{P(AB)}{P(B)} = \dfrac{25/300}{100/300} = \dfrac{25}{100} = 0{,}25$

Indica la probabilidad del hecho "La persona escogida es gerencial (A)" dado el hecho "La persona es universitaria (B)"

b- $P(B/A) = \dfrac{P(BA)}{P(A)} = \dfrac{25/300}{30/300} = \dfrac{25}{30} = 0{,}83$

Indica la probabilidad del hecho "La persona es universitaria (B)" dado el hecho "La persona escogida es gerencial (A)"

c- $P(A/D) = \dfrac{P(AD)}{P(D)} = \dfrac{5/300}{200/300} = \dfrac{5}{200} = 0{,}025$

Indica la probabilidad del hecho "La persona es Gerencial (A)" dado el hecho "La persona es no universitaria (D)".

d- $P(D/A) = \dfrac{P(DA)}{P(A)} = \dfrac{5/300}{30/300} = \dfrac{5}{30} = 0{,}17$

Indica la probabilidad del hecho "La persona es no universitaria (D)" dado el hecho "La persona es Gerencial (A)".

e- $P(C/B) = \dfrac{P(CB)}{P(B)} = \dfrac{75/300}{100/300} = \dfrac{75}{100} = 0{,}75$

Indica la probabilidad del hecho "La persona es no gerencial (C)" dado el hecho "La persona es universitaria (B)".

f- $P(B/C) = \dfrac{P(BC)}{P(C)} = \dfrac{75/300}{270/300} = \dfrac{75}{270} = 0{,}28$

Indica la probabilidad del hecho "La persona es universitaria (B)" dado el hecho "La persona es no gerencial (C)".

g- $P(C/D) = \dfrac{P(CD)}{P(D)} = \dfrac{195/300}{200/300} = \dfrac{195}{200} = 0{,}975$

Indica la probabilidad del hecho "La persona es no gerencial (C)" dado el hecho "La persona es no universitaria (D)".

h- $P(D/C) = \dfrac{P(DC)}{P(C)} = \dfrac{195/300}{270/300} = \dfrac{195}{270} = 0{,}72$

Indica la probabilidad del hecho "La persona es no universitaria (D)" dado el hecho "La persona es no gerencial (C)".

Obsérvese:

1) Que las probabilidades que se detallan a continuación, además de ser distintas, no son complementarias, es decir, que su suma no es igual a la unidad:

P(*A/B)*	y	P (*B/A*)
P(*A/D*)	y	P (*D/A*)
P(*C/B*)	y	P (*B/C*)
P(*C/D)*	y	P (*D/C*)

2) Que las probabilidades de los hechos que tienen idéntica condición son complementarias. Así:

$$P(A/B) + P(C/B) = \frac{25}{100} + \frac{75}{100} = 1$$

$$P(B/A) + P(D/A) = \frac{25}{30} + \frac{75}{30} = 1$$

$$P(A/D) + P(C/D) = \frac{5}{200} + \frac{195}{200} = 1$$

$$P(B/C) + P(D/C) = \frac{75}{270} + \frac{195}{270} = 1$$

Esto ocurre, pues la condición implica el espacio probabilístico reducido, y la probabilidad en él es 1.

2- Probabilidades Conjuntas

a- $P(A \cap B) = P(B \cap A) = P(A/B)\,P(B) = P(B/A)\,P(A)$

$$0{,}25 \times 0{,}33 = 0{,}83 \times 0{,}10$$
$$\mathbf{0{,}083 = 0{,}083}$$

Indica la probabilidad de ocurrencia conjunta de gerencial y universitario.

b- $P(A \cap D) = P(D \cap A) = P(A/D)\,P(D) = P(D/A)\,P(A)$

$$0{,}25 \times 0{,}67 = 0{,}17 \times 0{,}10$$
$$\mathbf{0{,}017 = 0{,}017}$$

Indica la probabilidad de ocurrencia conjunta de gerencial y no universitario.

c- $P(C \cap B) \; = P(B \cap C) = P(C / B)\; P(B) = P(B / C)\; P(C)$

$$0{,}75 \text{ x } 0{,}33 = 0{,}28 \text{ x } 0{,}90$$
$$\mathbf{0{,}25 = 0{,}25}$$

Indica la probabilidad de ocurrencia conjunta de no gerencial y universitario.

d- $P(C \cap D) \; = P(D \cap C) = P(C / D)\; P(D) = P(D / C)\; P(C)$

$$0{,}975 \text{ x } 0{,}67 = 0{,}72 \text{ x } 0{,}90$$
$$\mathbf{0{,}65 = 0{,}65}$$

Indica la probabilidad de ocurrencia conjunta de no gerencial y no universitario.

3- **Probabilidades Marginales**

a- $P(A) = P(A \cap B) \; + P(A \cap D) = 0{,}083 + 0{,}017 = 0{,}10$

Indica la probabilidad de que el empleado sea gerencial.

b- $P(C) = P(C \cap B) \; + P(C \cap D) = 0{,}25 + 0{,}65 = 0{,}90$

Indica la probabilidad de que el empleado sea no gerencial.

c- $P(B) = P(A \cap B) \; + P(C \cap B) = 0{,}083 + 0{,}25 = 0{,}333$

Indica la probabilidad de que el empleado sea un graduado universitario.

d- $P(D) = P(A \cap D) \; + P(C \cap D) = 0{,}17 + 0{,}65 = 0{,}667$

Indica la probabilidad de que el empleado sea un no universitario.

11. Independencia y Dependencia Estadística

En principio diremos que:

- Dos eventos son dependientes si la probabilidad de ocurrencia de uno es afectada por la ocurrencia del otro.

- Dos eventos son independientes si la probabilidad de ocurrencia de uno no es afectada por la ocurrencia del otro.

La dependencia e independencia estadística se relacionan con la naturaleza del proceso aleatorio de selección de los elementos, o sea con el proceso de muestreo.

En general, los *hechos dependientes* son generados por muestreo aleatorio *sin reposición* y, los *hechos independientes* son generados por muestreo aleatorio *con reposición.*

El muestreo aleatorio sin reposición es un proceso de selección al azar de **n** unidades, que constituyen la muestra, de una población de **N** unidades, sin devolver a la población ninguna unidad escogida antes de extraer otra.

El muestreo aleatorio con reposición es un proceso de selección al azar de **n** unidades, que constituyen la muestra, de una población de **N** unidades, donde cada unidad extraída es reintegrada a la población antes de extraer otra.

<u>Ejemplo:</u>

Supongamos que se elige una bolilla al azar de una bolsa que posee cuatro bolillas azules y seis blancas. Se observa el color y

1- Se repone en la bolsa

2- No se repone en la bolsa

Se pregunta: ¿Cuál es la probabilidad de que la segunda bolilla sea azul, en ambos casos?

Entonces, tenemos: 10 bolillas: 4 azules
6 blancas

A partir de ello, definimos: A = (Bolillas azules)
B = (Bolillas blancas)

Veamos qué ocurre:

1- Si se repone en la bolsa la primer bolilla extraída, tenemos:

<u>1° Elección</u>:

P(A) = 4/10 = 0,40
P (B) = 6/10 = 0,60

<u>2º Elección</u>:

P(A) = 4/10 = 0,40, puesto que la respuesta, no depende del resultado de la primera elección, pues la bolilla fue sustituida, entonces los dos sucesos son independientes, a saber, el resultado de una no modifica el resultado de la otra, en términos de no afectar la probabilidad en la segunda extracción.

2- Si no se repone la primera bolilla, el resultado dependerá de lo ocurrido en la primera extracción.

<u>1° Elección</u>:

P(A) = 4/10 = 0,40
P (B) = 6/10 = 0,60

2° Elección:

2.1. Si la primer bolilla sacada fue azul, quedan 9 bolillas en la bolsa, de las cuales 3 son azules, entonces P(*A/A*) = 3/9 (Probabilidad de que la segunda bolilla sea azul, dado que la primer bolilla es azul).

2.2. Si la primer bolilla sacada fue blanca, quedan 9 bolillas en la bolsa, de las cuales 4 son azules, entonces P(*A/B*) = 4/9 (Probabilidad de que la segunda bolilla sea azul, dado que la primer bolilla es blanca).

Una vez realizada la distinción entre eventos dependientes e independientes, diremos:

La regla que rige la probabilidad de ocurrencia conjunta de dos hechos que son independientes es un caso especial de la Ley Multiplicatoria General. Esta regla especial la da el siguiente Teorema:

"Dos hechos, *A* y *B*, definidos en el mismo espacio de muestra, se dice que son independientes sí, y solo sí, la probabilidad de la ocurrencia conjunta de *A* y *B* es igual al producto de sus respectivas probabilidades individuales".

Es decir, *A* y *B* son hechos independientes sí, y solo sí

$$P(A \cap B) = P(A) \, x \, P(B)$$

Recuérdese que: $P(A \cap B) = P(A/B)\,P(B)$ y $P(B \cap A) = P(B/A)\,P(A)$, pero si los eventos son independientes

$$P(A/B) = P(A) \text{ Y } P(B/A) = P(B)$$

Siempre que: $P(A) \neq 0$ y $P(B) \neq 0$.

Es decir que la probabilidad de uno no está condicionada a la ocurrencia del otro.

Además dijimos que:

$$P(A \cap B) = P(B \cap A)$$

Reemplazando por sus iguales:

$$P(A/B)\,P(B) = P(B/A)\,P(A),$$

y si los eventos son independientes:

$$P(A)\,P(B) = P(B)\,P(A)$$

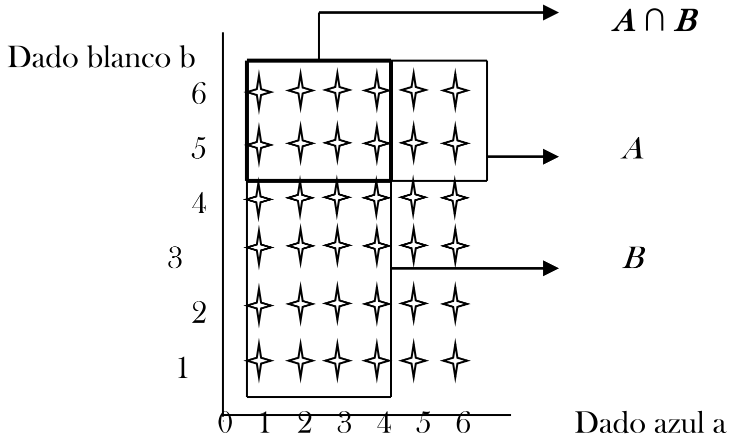
La función de independencia estadística conduce a una definición formal de dependencia estadística: Dos hechos, *A* y *B* definidos en el mismo espacio de muestra, se dice que son dependientes sí, y solo sí:

$$P(A \cap B) \neq P(A)\, P(B)$$
$$P(A \cap B) = P(A/B)\, P(B) = P(B/A)\, P(A)$$

Ejemplo:

Supongamos que se tira un dado blanco, b, y un dado azul, a, ¿Cuál es la probabilidad de que b ≥ 5 y a ≤ 4?

El espacio muestral, consta de 36 combinaciones igualmente posibles del dado blanco y del dado azul, cada una de ellas con seis resultados posibles, y que el hecho que se considera requiere que sean satisfechos simultáneamente las dos condiciones.



$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{8/36}{24/36} = \frac{8x36}{24x36} = \frac{288}{864} = \frac{12}{36} = P(A)$$

Si *A* es el hecho de que b ≥ 5 y a ≤ 4, debemos conocer entonces el número de puntos de muestral que ambos conjuntos tienen en común, es decir *A* ∩ *B*.
En la figura podemos ver que hay ocho puntos en el conjunto de intersecciones, entonces:

$$P(A \cap B) = \frac{8}{36}$$

Además, hallamos que hay doce puntos de muestra en *A* y veinticuatro en ***B***, entonces:

$$P(A) = \frac{12}{36} \quad y \quad P(B) = \frac{24}{36}$$

$$P(A \cap B) = P(A)\, P(B) = \frac{12}{36}\, x\, \frac{24}{36} = \frac{288}{1296} = \frac{8}{36}$$

Igual que antes, por lo tanto, concluimos que A y B son eventos independientes.

La extensión de la definición de independencia a tres ó más hechos, no es inmediata. E_1, E_2, E_3,..., E_n, son independientes sí, y solo sí, son independientes por pares, por triples, por cuádruplos, etc.

En particular, tres hechos A, B y C son completamente independientes sí y solo sí:

$$P(A \cap B) = P(A)\,P(B)$$
$$P(A \cap C) = P(A)\,P(C)$$
$$P(B \cap C) = P(B)\,P(C) \quad \text{Y}$$
$$P(A \cap B \cap C) = P(A)\,P(B)\,P(C)$$

Por lo tanto, los hechos deben ser independientes no sólo por pares, sino también por triples.

Clarificaremos ahora la diferencia entre eventos mutuamente excluyentes y eventos independientes. Si lo considerado es la unión de los eventos, nos interesaremos en saber si son eventos mutuamente excluyentes. Si lo considerado es la intersección de eventos, nos interesaremos por la independencia de eventos.
Además, si dos eventos son mutuamente excluyentes y si ningún evento tiene probabilidad 0, los dos eventos deben ser estadísticamente dependientes.

Sin embargo, si dos eventos son dependientes, no se deduce necesariamente que sean mutuamente excluyentes.

Por último, si A y B son independientes y si P (A) ≠ 0 y P (B) ≠ 0, entonces A y B no pueden ser mutuamente excluyentes, porque para que dos eventos sean independientes la probabilidad conjunta debe ser el producto de las probabilidades marginales o individuales. En este caso la $P(A \cap B)=0$, ya que $A \cap B = \Phi$ y la $P(A) \times P(B)$ es mayor a cero.

12. Teorema o Regla De Bayes

Todos los ejemplos presentados hasta el momento contaban con la probabilidad de los eventos calculados "a priori". Así, cuando tiramos una moneda balanceada, decimos a priori que la probabilidad de salir cara en una tirada es de 1/2 y la de sello también 1/2, o sea que establecemos el espacio probabilístico y sus correspondientes eventos con las probabilidades asociadas antes de realizar la prueba.

En ciertos casos es posible establecer algunas características o determinados hechos, vinculados con eventos cuya probabilidad a priori es conocida, los que una vez ocurridos pueden ser analizados con el fin de determinar si los mismos han ocurrido conforme a determinadas causas. Estas probabilidades se conocen con el nombre de "probabilidades a posteriori" y conforma una regla de probabilidades muy útil denominada *Teorema De Bayes o Regla De Bayes.*

Este Teorema ha ganado mucho auge y no es más que una fórmula para calcular probabilidades condicionales. Dada una tabla de probabilidades compuestas, es sencillo calcular probabilidades condicionales. Solo hay que dividir una probabilidad compuesta por la probabilidad marginal, que es el total de las probabilidades compuestas en una fila o columna dadas, según sea el suceso que se supone dado. Si no se dispone de tabla de probabilidades compuesta, entonces hay que establecer la tabla.

Un breve ejemplo ilustrará el proceso.

Supóngase que en una habitación oscura se colocan tres bolsas.

En la bolsa 1 (B_1), hay seis bolillas rojas (R) y cuatro blancas (B); en la bolsa 2 (B_2), hay siete bolillas rojas y tres blancas; en la bolsa 3 (B_3), hay cinco rojas y cinco blancas. Si alguien va a la habitación y toma una bolsa al azar de la cual saca una bolilla roja, ¿Cuál es la probabilidad de que la bolilla roja se halla tomado de B_1?. De B_2?. De B_3?.

Suponiendo idénticas las bolsas, la probabilidad de elegir una cualquiera de las tres es 1/3, que es la probabilidad marginal.

El problema exige el cálculo de tres probabilidades condicionales, $P(B_1/R)$, $P(B_2/R)$, $P(B_3/R)$.

Para esto hay que establecer una tabla de probabilidades compuestas:

	R	B	Total
B_1	$6/30=0{,}20$	$4/30=0{,}13$	$10/30=1/3=0{,}33$
B_2	$7/30=0{,}23$	$3/30=0{,}10$	$10/30=1/3=0{,}33$
B_3	$5/30=0{,}17$	$5/30=0{,}17$	$10/30=1/3=0{,}33$
Total	$18/30=0{,}60$	$12/30=0{,}40$	1

Entonces:

$$P(B_1/R)=\frac{P(B_1\ R)}{P(R)}=\frac{\frac{6}{30}}{\frac{18}{30}}=\frac{6}{18}=\underline{0{,}33} \qquad (1)$$

$$P(B_2/R)=\frac{P(B_2\ R)}{P(R)}=\frac{\frac{7}{30}}{\frac{18}{30}}=\frac{7}{18}=\underline{0{,}39} \qquad (2)$$

$$P(B_3/R)=\frac{P(B_3\ R)}{P(R)}=\frac{\frac{5}{30}}{\frac{18}{30}}=\frac{5}{18}=\underline{0{,}28} \qquad (3)$$

Recordemos:

1- $P(A/B)=\dfrac{P(A\cap B)}{P(B)}$ y $P(B/A)=\dfrac{P(B\cap A)}{P(A)}$

Luego: $P(A \cap B) = P(B \cap A) = P(A/B)\,P(B) = P(B/A)\,P(A)$

2- Toda probabilidad marginal ó individual es suma de probabilidades conjuntas.

En base a esto, reemplazamos en (1), (2) y (3):

$$P(B_1/R)=\frac{P(B_1\ R)}{P(R)}=\frac{P(R/B_1)\ P(B_1)}{P(RB_1)+P(RB_2)+P(RB_3)}=\frac{P(R/B_1)\ P(B_1)}{P(R/B_1)\ P(B_1)+P(R/B_2)\ P(B_2)+P(R/B_3)\ P(B_3)}$$

$$P(B_2/R)=\frac{P(B_2\ R)}{P(R)}=\frac{P(R/B_2)\ P(B_2)}{P(RB_1)+P(RB_2)+P(RB_3)}=\frac{P(R/B_2)\ P(B_2)}{P(R/B_1)\ P(B_1)+P(R/B_2)\ P(B_2)+P(R/B_3)\ P(B_3)}$$

$$P(B_3/R)=\frac{P(B_3\ R)}{P(R)}=\frac{P(R/B_3)\ P(B_3)}{P(RB_1)+P(RB_2)+P(RB_3)}=\frac{P(R/B_3)\ P(B_3)}{P(R/B_1)\ P(B_1)+P(R/B_2)\ P(B_2)+P(R/B_3)\ P(B_3)}$$

En general, si hay n ***B*** y se desea calcular la probabilidad condicional **P** **(*B_i/R*)** para i = 1,2,3...,n, la fórmula de Bayes es:

$$P(B_i/R)=\frac{P(B_i\ R)}{P(R)}=\frac{P(R/B_i)\ P(B_i)}{P(RB_1)+P(RB_2)+\ldots+P(RB_n)}=\frac{P(R/B_i)\ P(B_i)}{P(R/B_1)\ P(B_1)+P(R/B_2)\ P(B_2)+\ldots+P(R/B_n)\ P(B_n)}$$

Siendo las **B**, sucesos de cierto tipo, y **R** un suceso de otro tipo, ú observación.

Así, para el ejemplo, tenemos:

$$P(B_1/R)=\frac{P(R/B_1)\ P(B_1)}{P(R/B_1)\ P(B_1)+P(R/B_2)\ P(B_2)+P(R/B_3)\ P(B_3)}$$

Calculamos:

$$P(R/B_1)=\frac{P(RB_1)}{P(B_1)}=\frac{\frac{6}{30}}{\frac{10}{30}}=\frac{6}{10}=\underline{0{,}60}$$

$$P(R/B_2)=\frac{P(RB_2)}{P(B_2)}=\frac{\frac{7}{30}}{\frac{10}{30}}=\frac{7}{10}=\underline{0{,}70}$$

$$P(R/B_3)=\frac{P(RB_3)}{P(B_3)}=\frac{\frac{5}{30}}{\frac{10}{30}}=\frac{5}{10}=\underline{0{,}50}$$

Reemplazando:

$$P(B_1/R) = \frac{\frac{10}{30} x \frac{6}{10}}{\frac{10}{30} x \frac{6}{10} + \frac{10}{30} x \frac{7}{10} + \frac{10}{30} x \frac{5}{10}} = \frac{\frac{60}{300}}{\frac{60}{300} + \frac{70}{300} + \frac{50}{300}} = \frac{\frac{6}{30}}{\frac{6}{30} + \frac{7}{30} + \frac{5}{30}} = \frac{\frac{6}{30}}{\frac{18}{30}} = \frac{6}{18}$$

Igual que antes.

Obsérvese:

1- Que el numerador es P (B_1R) y el denominador es P (B_1R) + P (B_2R) + P (B_3R) = P(R), según la tabla de probabilidades compuestas construida.

Así pues, la fórmula de Bayes está concebida para calcular probabilidades condicionales marchando a la inversa, o sea si se conoce P(R/B$_i$) se calcula la P(B$_i$/R), que es la probabilidad de B$_i$ una vez hecha la observación, por ello, es una probabilidad posterior o a posteriori, en tanto que la probabilidad de B, o sea P(B$_i$), se llama probabilidad anterior ó a priori, indicando que es la probabilidad antes de hacer ninguna observación.

Trataremos a continuación otro ejemplo:

Supóngase que un fabricante tiene tres máquinas, que se llamarán X, Y, Z, que producen partes idénticas. Se sabe que el 5% de las partes producidas por la máquina X son defectuosas (D), y que los porcentajes de partes defectuosas que producen Y y Z son el 10 y el 15 por ciento respectivamente. Se mezclan los productos de las tres máquinas y no hay manera de reconocer cual ha sido producido por cual máquina. Pero, se sabe, que las tres máquinas tienen igual capacidad y que funcionan al mismo ritmo de producción. Si se toma una parte producida, al azar, y se la encuentra defectuosa, ¿Cuál es la probabilidad de que provenga de la máquina Y? . Es decir, se pide calcular P(Y/D).

P(X) = 1/3 = 0,33 P (Y) = 1/3 = 0,33 P (Z) = 1/3 = 0,34

P (D/X) = 0,05 P (D/Y) = 0,10 P (D/Z) = 0,15

Entonces, con los datos que contamos, aplicamos:

$$P(Y/D) = \frac{P(D/Y)P(Y)}{P(D/X)P(X) + P(D/Y)P(Y) + P(D/Z)P(Z)} =$$

$$= \frac{0{,}10 X\, 0{,}33}{(0{,}05 X\, 0{,}33) + (0{,}10 X\, 0{,}33) + (0{,}15 X\, 0{,}34)} = \frac{0{,}033}{0{,}0165 + 0{,}033 + 0{,}0495} = \frac{0{,}033}{0{,}099} = \underline{0{,}33}$$

En realidad, el mérito de Bayes no fue tanto la originalidad sino expresar la probabilidad condicional en función de la probabilidad de la intersección. Además, el honor del teorema que lleva su nombre no es completamente suyo, ya que él no estaba en condiciones de formular con probabilidades totales. Fue Pierre-Simon Laplace (1749-1827) quien desarrolló la mayor parte del Teorema de Bayes.

Laplace aplicó el teorema a problemas de la mecánica celeste, la estadística médica e, incluso, a la jurisprudencia.

13. Aplicaciones de la Teoría de Probabilidad: Procesos Estocásticos

La probabilidad clásica sólo se ocupa de problemas estacionarios en el tiempo, mientras que físicos, biólogos e ingenieros están interesados en procesos que evolucionan en el tiempo. Algunos ejemplos de estos problemas son el número y la duración de las llamadas telefónicas, sistemas dinámicos de población, movimiento de partículas que chocan entre ellas, difusión entre líquidos, decaimiento radioactivo.

El concepto de procesos estocásticos, basado en la definición axiomática de la probabilidad, es el siguiente: Sea Ω el conjunto de sucesos elementales y t un parámetro continuo. Un proceso estocástico es la función de dos argumentos $\xi(t) = \Phi(\omega, t)$, $\omega \in \Omega$; para cada valor de t, la función $\Phi(\omega, t)$ es sólo una función de ω y, por consiguiente, una variable aleatoria; para cada valor fijo de ω, la función $\Phi(\omega, t)$ depende sólo de t y es una función de una variable real, función que recibe el nombre de "*realización del proceso estocástico*". Un proceso aleatorio puede verse como la colección de variables aleatorias $\xi(t)$ que dependen del parámetro t o como la colección de las realizaciones del proceso $\xi(t)$. Es natural asignar una medida de probabilidad en el espacio de funciones de las realizaciones para definir el proceso.

Los primeros en estudiar estos procesos fueron físicos como Max Planck y Fokker, y matemáticos como Markov o Slutsky, pero no fue hasta la década de 1930 cuando Kolmogorov y Khinchine construyeron la teoría general y rigurosa de los procesos estocásticos.

Resumen

<table>
<tr><th>REGLAS PARA LA PROBABILIDAD DE LA UNIÓN DE LOS EVENTOS</th><th>REGLAS PARA LA PROBABILIDAD DE LA INTERSECCIÓN DE LOS EVENTOS</th></tr>
<tr><td>Eventos Mutuamente Excluyentes
$(A_i \cap A_j \ = \emptyset \quad i \neq j = 1,2,\ldots)$

$P(A \cup B)) = \ P(A) + P(B)$</td><td>Eventos Dependientes

$P(A \cap B) \neq \ P(A) \ \ P(B)$

$P(B \cap A) \neq \ P(B) \ \ P(A)$

Para
P (A) ≠ 0 y P (B) ≠ 0</td></tr>
<tr><td rowspan="2">Eventos No Mutuamente Excluyentes
$(A_i \cap A_j \ \neq \emptyset)$

$P(A \cup B)) = \ P(A) + P(B) - P(A \cap B)$</td><td>Eventos Independientes

$P(A \cap B) = \ P(A) \ \ P(B)$

$P(B \cap A) = \ P(B) \ \ P(A)$</td></tr>
<tr><td>Eventos Dependientes

$P(A \cap B) = \ P(A / B) \ \ P(B)$

$P(B \cap A) = \ P(B / A) \ \ P(A)$</td></tr>
</table>

LEY DE PROBABILIDAD CONDICIONAL	$P(A/B) \ = \dfrac{P(A \cap B)}{P(B)} \quad ; \quad P(B) \rangle 0$ $P(B/A) \ = \dfrac{P(A \cap B)}{P(A)} \quad ; \quad P(A) \rangle 0$
REGLA DE BAYES	$P(B_i / R) = \dfrac{P(R/B_i) \ P(B_i)}{P(R/B_1) \ P(B_1) + P(R/B_2) \ P(B_2) + \ldots + P(R/B_n) \ P(B_n)}$

CAPITULO Nº 5: Variables Aleatorias y Distribuciones De Probabilidad

Objetivos Específicos

Que el estudiante:

Comprenda *la diferencia entre variable y variable aleatoria.*

Distinga *cuándo utiliza distribuciones de frecuencias y de probabilidad.*

Calcule *los parámetros en las distribuciones de probabilidad.*

Interprete *la esperanza matemática y la desviación estándar.*

Comprenda *la utilidad de los momentos.*

Contenidos

1. Variable aleatoria.
- ***3.1.*** *Generalidades.*
- ***3.2.*** *Definición.*
- ***3.3.*** *Variables aleatorias discretas y continuas.*

2. Distribuciones de probabilidad.
- ***2.1.*** *Función de probabilidad para variables discretas: función de cuantía.*
- ***2.2.*** *Función de distribución para variables aleatorias discretas.*
- ***2.3.*** *Función de densidad y función de distribución para variables aleatorias continuas.*

3. Los parámetros en las distribuciones de probabilidad.
- ***3.1.*** *Esperanza matemática. Propiedades.*
- ***3.2.*** *Varianza. Propiedades.*
- ***3.3.*** *Desviación estándar.*

4. Momentos en las distribuciones de probabilidades: *momento natural de orden k y momento centrado de orden k*

1. Variable Aleatoria

1.1. Generalidades

Hemos dado anteriormente el concepto de espacio probabilístico, definiéndolo como el conjunto de resultados posibles de un experimento aleatorio, (que pueden ser diferentes en la repetición de las pruebas, no obstante que las mismas hayan sido realizadas en idénticas condiciones). Luego a cada resultado posible, le asociamos su correspondiente probabilidad, pues cada uno de ellos constituye un evento elemental o simple.

Ahora bien, tales resultados pueden ser hechos, sucesos, etc., de distinta naturaleza que dependen del problema analizado y de su definición.

En algunos casos, los resultados posibles pueden ser presencia de números, como en la tirada de un dado, donde el evento suele considerarse el número inscripto en la cara superior; o cuando se controla el número de errores cometidos en un libro de contabilidad.

En otros casos, el fenómeno aleatorio no da como resultado un número, tal es el caso de la tirada de una moneda, donde los resultados son cara o sello; en procesos productivos, pueden ser presencia de artículo bueno o presencia de artículo defectuoso.

A los fines de posibilitar aún más el manejo del análisis matemático, siempre es posible representar a los resultados, cualquiera sean los hechos o conceptos que representen, mediante un sistema o conjunto numérico, de tal forma que podamos identificar a cada uno de ellos por medio de un número real o un intervalo numérico.

Es decir que a cada uno de los resultados posibles de un fenómeno aleatorio (a cada evento simple) se le asocia un número real de tal manera que siempre podamos expresar numéricamente a los resultados de un experimento aleatorio.

Este nuevo conjunto numérico que representará al conjunto de resultados posibles o eventos elementales, será entonces el conjunto de valores que puede asumir una nueva variable, que llamaremos *Variable Aleatoria* (casual, estocástica, ó variable de azar) y cada valor, o conjunto de valores, intervalos, unión o intersección de intervalos de dicha variable, tendrán asociados una probabilidad de presentación, que es la probabilidad del evento que los valores de dicha variable representan.

Simbolizaremos la variable aleatoria por: x, y, etc.

Ejemplo:

1- Clasificación de los alumnos en un examen final. Los resultados del examen pueden expresarse por grados, como insuficiente, suficiente, bueno, distinguido, sobresaliente, o cualquier otra expresión, que permita medir lo que sabe el alumno que ha rendido. Se puede establecer, por ejemplo la siguiente regla:

$$x = \left| \begin{array}{ll} 0 & \textit{Reprobado} \\ 1 & \textit{Suficiente} \\ 2 & \textit{Bueno} \\ 3 & \textit{Distinguido} \\ 4 & \textit{Sobresaliente} \end{array} \right.$$

2- Una población dicotomizada clasificada en varones y mujeres. La variable x asociada al experimento aleatorio “observación de un individuo de esta población”, puede ser:

$$x = \left| \begin{array}{ll} 1 & \textit{si el individuo es varón} \\ 0 & \textit{si el individuo es mujer} \end{array} \right.$$

3- Si en la tirada de una moneda decimos que los eventos son “salida de cara” y “salida de sello”, podemos reemplazar o representar a cada uno de estos eventos mediante un número.

Por ejemplo, al evento “salida de cara” lo representamos con el número 1 y al evento “salida de sello” con el número 0, que corresponderá a una variable aleatoria x.

De esta manera, si x asume el valor 1, implicará que se ha presentado el evento “salida de cara” y si x asume el valor 0, implicará que se ha presentado el evento “salida de sello”.

Simbólicamente:

Si (c), es el evento “salida de cara”, con $P(c) = \frac{1}{2}$, asociada.

Y (s), es el evento “salida de sello”, con $P(s) = \frac{1}{2}$, asociada.

Definimos la variable aleatoria X, tal que:

$$\text{Si } (x = 1) \Rightarrow (c) \text{ y } \therefore\ P(c) = P(x = 1) = \frac{1}{2} \text{ y}$$

$$\text{Si } (x = 0) \Rightarrow (s) \text{ y } \therefore\ P(s) = P(x = 0) = \frac{1}{2}$$

Es decir, que los eventos que antes se representaban con la simbología de conjuntos, se representan ahora mediante números o intervalos numéricos que corresponden a una variable aleatoria donde sus probabilidades de presentación, son las probabilidades de presentación de los resultados o eventos que representan.

O sea, decir “la probabilidad de salida de cara” es igual a 0,50, equivale a decir, “la probabilidad de que la variable aleatoria asuma el valor 1” es igual a 0,50. Por lo tanto la probabilidad de que la variable aleatoria asuma cualquier valor distinto de 0 y 1 es igual a 0, pues es el evento imposible, y la probabilidad de que asuma el valor 0 ó 1 es igual a 1, pues ello implica la probabilidad del espacio (Ω), pues

Ω = (c*, s),* es con la variable aleatoria:

Ω = (0, 1)

<u>Ejemplo</u>:

En la tirada de un dado, la variable aleatoria puede asumir seis valores:

x_1=1, x_2=2, x_3=3, x_4=4, x_5 =5, x_6 =6, y cada uno de dichos valores representan los eventos:

"salida del Nº 1", "salida del N° 2",..., "salida del N° 6".

Si consideramos ahora el evento: A_1 = (salida del Nº par), podemos representarlo por un conjunto de valores de la variable aleatoria X, de la siguiente manera:

A_1 = ($1 \leq x \leq 6$); para X par

De igual forma, si definimos al evento A_2 = (salida del Nº menor ó igual a tres), en términos de variable aleatoria, nos quedará:

A_2 = ($1 \leq x \leq 3$); para x natural

Es importante aclarar que si un determinado evento se representa mediante un conjunto de números reales, o un intervalo numérico, ello no significa que la variable aleatoria deba asumir todos los valores numéricos. Basta que asuma un solo valor para que ello implique la presencia del evento representado por el conjunto numérico o intervalo. Esta apreciación es coincidente con lo visto en la unidad anterior en relación a que un evento ha ocurrido si al menos uno de sus elementos se ha presentado.

<u>Ejemplo</u>:

En la salida de número par, el evento está representado por el conjunto de números:

A_1 = (2, 4, 6)

Que mediante la variable aleatoria queda representada por

A_1 = ($1 \leq x \leq 6$) para x par

Si la variable aleatoria asume el valor $\begin{cases} x = 2 \\ \text{ó} \\ x = 4 \\ \text{ó} \\ x = 6 \end{cases}$

Cualquiera de estos tres valores que tome, implicará la presencia del evento A_1 (Nº par).

1.2. Definición

En función de lo analizado anteriormente, podemos expresar que una variable aleatoria es una función real valorada, definida sobre los eventos elementales de un espacio probabilístico, donde a cada evento elemental le corresponde un número que es el valor que asume la variable aleatoria para dicho evento elemental. Una variable es aleatoria si puede asumir cualquier valor en un cierto recorrido con una cierta probabilidad.

1.3. Variables Aleatorias Discretas y Continuas

Como vimos en la Unidad 1, las Variables pueden ser Discretas o Continuas. Luego, las *Variables Aleatorias,* también puede clasificarse de esta misma forma.
Así, tendremos:

Variables Aleatorias Discretas

Cuando tienen una cantidad numerable de valores posibles, que podrán ser un número finito de valores, ó un número infinito pero numerable de valores reales.

$X_1, X_2, X_3, \ldots, X_N$ Ó bien $X_1, X_2, X_3, \ldots\ldots, X_\infty$

Variables Aleatorias Continuas

Cuando pueden asumir cualquier valor real, es decir los infinitos valores del conjunto de los números reales, o cuando pueden asumir cualquier valor dentro de un determinado intervalo, de tal manera que entre uno y otro valor pueden diferir hasta en un infinitésimo.

2. Distribuciones de Probabilidad

Se ha dicho que los que tienen probabilidad son los eventos y que cualquier sistema completo de eventos constituyen el espacio muestral cuya probabilidad es igual a 1.

La probabilidad de los eventos se traslada a los valores de las variables aleatorias, y cuando se dice probabilidad de que la variable aleatoria tome un cierto valor, implica decir probabilidad que ocurra alguno de aquellos eventos simples de los cuales depende ese valor de la variable aleatoria.

Entonces una Distribución de Probabilidad muestra a través de una tabla, gráfico ó fórmula, todos los valores posibles que puede asumir una variable aleatoria y su correspondiente probabilidad de presentación.

Para una mayor claridad se distingue entre:

1- Funciones de Probabilidad, que reciben el nombre de:
   1. a. Función de Cuantía, si la variable es Discreta.
   1. b. Función de Densidad, si la variable es Continua.

2- Funciones de Distribución ó Funciones de Acumulación.

2.1. Función De Probabilidad Para Variables Aleatorias Discretas: Función de Cuantía

Cuando se cuenta con todos los posibles valores de una variable aleatoria y la correspondiente probabilidad de cada uno de esos valores, se tiene la función de probabilidad de una variable aleatoria, que permite obtener la probabilidad para cada uno de los valores de la variable correspondiente a un determinado experimento.

En general, $P(x = x_i)$, indica la probabilidad de que la variable aleatoria x asuma un valor real x_i, lo cual implica la probabilidad de que se presente un determinado evento, representado por ese valor x_i de la variable aleatoria.

Y las probabilidades asociadas se simbolizan mediante la expresión:

$$P\left(x = x_i \right) \left| \begin{array}{c} i = 1,2,3,\ldots,N \\ ó \\ i = 1,2,3,\ldots,\infty \end{array} \right.$$

Por comodidad se suele representar:

$$P\left(x = x_i \right) = \left| \begin{array}{ll} P_i & \left| \begin{array}{l} i = 1,2,3,\ldots,N \\ i = 1,2,3,\ldots,\infty \end{array} \right. \\ 0\ \forall \text{ otro } i & \end{array} \right. \quad \text{siendo} \quad \sum_i P_i = 1 \quad \text{para } P_i \geq 0$$

<u>Ejemplo:</u>

Si se considera la tirada de un dado, entonces la función de cuantía es:

$x = x_i$	$P(x = x_i) = P_i$
1	$P(x = 1) = 1/6$
2	$P(x = 2) = 1/6$
3	$P(x = 3) = 1/6$
4	$P(x = 4) = 1/6$
5	$P(x = 5) = 1/6$
6	$P(x = 6) = 1/6$
	$\sum_i P_i = 1$

Esta función de cuantía queda definida mediante la siguiente expresión:

$$P\left(x = x_i \right) \left| \begin{array}{ll} 1/6; & \text{para } x = 1,2,3,4,5,6 \\ 0; & \forall \text{ otro } x \end{array} \right.$$

Y responde a la expresión genérica:

$$P\left(x = x_i\right) = \begin{vmatrix} P_i; & \text{para } x = 1,2,3,4,5,6 \\ 0; & \forall \text{ otro } x \end{vmatrix}$$

Donde P_i es una constante.

De acuerdo a lo visto, al tratarse de una variable aleatoria discreta, las probabilidades tendrán algún valor en todos los puntos o valores reales que asuma x, donde se verifica la presencia de un evento y será igual a 0 para todos los demás valores reales de x que no tengan asociados ó no representen eventos.

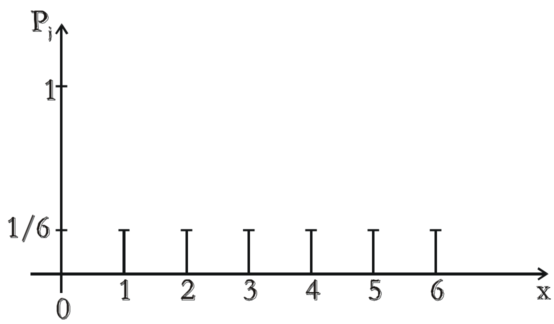
Los valores de la función de probabilidad corresponden a un sistema completo de eventos y por lo tanto su suma es igual a 1. En consecuencia la función de probabilidad de una variable aleatoria discreta (función de cuantía), debe satisfacer las siguientes condiciones:

a) $0 \leq P_i \leq 1 \qquad \forall\ x_i$

b) $\sum_{i=0}^{\infty} P_i = 1$ ó $\sum_{i=0}^{N} P_i = 1$

Esta función indica el valor de la probabilidad en el punto $x = x_i$, por eso es también llamada Probabilidad Puntual.

Representación Gráfica:



El caso anterior corresponde a una distribución uniforme, pero se puede tener una distribución de probabilidad no necesariamente uniforme.

<u>Ejemplo:</u>

Si se arrojan tres monedas; asignando el valor 1 al hecho de presentarse cara y 0 al de presentarse sello y nos interesa establecer el número de caras que se obtienen en la tirada.

Se pregunta: ¿Cuál es la función de probabilidad correspondiente al 1 en el recorrido de estos números?

Antes que nada se construye la tabla con los eventos elementales del espacio probabilístico, el número de caras y la probabilidad asociada a cada evento elemental.

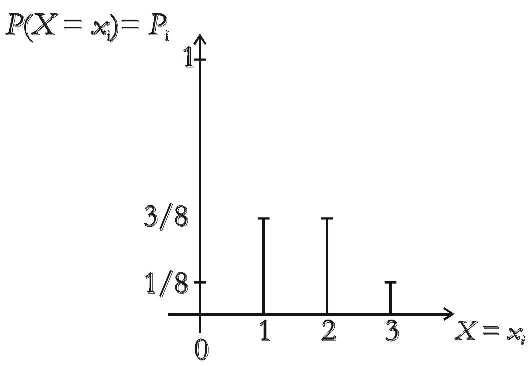
Evento Elemental	Nº de caras	Probabilidad
s,s,s	0	1/8
c,s,s	1	1/8
s,c,s	1	1/8
s,s,c	1	1/8
c,c,s	2	1/8
c,s,c	2	1/8
s,c,c	2	1/8
c,c,c	3	1/8

En base a esta información se determina la función de cuantía:

Distintos valores de $x = x_i$	$P(x = x_i) = P_i$
0	1/8
1	3/8
2	3/8
3	1/8

En este caso no se puede determinar una expresión general, pues P_i no es una constante.

Gráficamente:



2.2. Función de Distribución Para Variables Aleatorias Discretas

Sea, $P(x \leq x_i)$, donde x es un número real. $P(x \leq x_i)$ es la probabilidad de que la variable aleatoria x asuma valores menores ó iguales a x_i.

Esta probabilidad que es la suma de las probabilidades de todos los valores del recorrido menores o iguales a x, depende, en realidad, de este valor real x y se denota mediante $F(x)$, luego $F(x) = P(x \leq x_i) = \sum_{x \leq x_i} P_{x_i}$

$F(x)$ se llama *Función De Distribución o Función de Acumulación De La Variable Aleatoria x.* Esta función no es una probabilidad puntual, sino la probabilidad de un intervalo, mejor dicho de todos los eventos en el intervalo definido por $x \leq x_i$.

Ejemplo:

En la tirada de un dado definimos $P(x = x_i)$ y $P(x \leq x_i)$

$x = x_i$	$P(x = x_i) = P_i = P(x_i)$	$F(x_i) = \Pr(x \leq x_i)$
1	1/6	1/6
2	1/6	2/6
3	1/6	3/6
4	1/6	4/6
5	1/6	5/6
6	1/6	6/6 =1

Es decir:

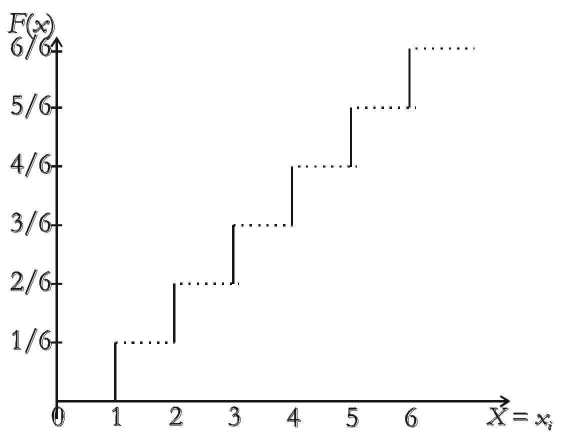
$$Pr\ (x \leq 1) = 1/6$$

$$Pr\ (x \leq 2) = Pr\ (x = 1) + Pr\ (x = 2) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6}$$

$$\vdots$$

$$Pr\ (x \leq 6) = Pr\ (x = 1) + \ldots + Pr\ (x = 6) = \frac{1}{6} + \frac{1}{6} + \ldots + \frac{1}{6} = \frac{6}{6} = 1$$

Gráficamente la *Función De Distribución* de una variable aleatoria discreta presenta "saltos" ó puntos de discontinuidad por la izquierda cuando $x = x_i$. Los saltos se presentan siempre que $p\ (x_i) \neq 0$ ya que estos saltos indican la medida de $p\ (x_i)$.



NOTA: Si se pidiera que se presente un nº≤ 3, ya no se estaría en presencia de un hecho simple, sino que el evento esperado puede presentarse de tres formas mutuamente excluyentes. Es decir, que si se tira el dado y sale cualquiera de los tres números: 1, 2, ó 3, se verifica el evento cuya probabilidad se solicita.

Una *Función De Distribución* para Variables Aleatorias Discretas satisface las siguientes condiciones:

1) *F (x)* es una función monótona creciente (no decreciente), continua a la derecha de cada punto, pues es acumulación de probabilidades que son no negativas.

2) Relaciones

$F(0) = Pr(x \leq 0) = 0$ *No hay valores p (x_i) para valores de $x < 0$*
$F(n) = Pr(x \leq n) = 1$ *Considerando que el total de valores posibles es n*

3) Dados dos números positivos enteros *a* y *b,* tales que $a < b$

$$P(a \leq x \leq b) = P(x \leq b) - P(x < a) = P(x \leq b) - P(x \leq a-1)$$

4) Dada una variable aleatoria *x* y un número positivo entero *a*

$$P(x \geq a) = 1 - P(x < a) = 1 - P(x \leq a-1)$$

2.3. Función de Densidad y Función de Distribución Para Variables Aleatorias Continuas

Muchas veces por las limitaciones propias de nuestros sentidos, ó por costumbre, se mencionan intervalos como si fueran puntos. Así, por ejemplo, cuando se menciona que una persona tiene 20 años no quiere decir que tiene 20 años exactamente, sino que ya ha cumplido 20 años pero aún no ha llegado a los 21.

Lo mismo resulta cuando se habla de estatura y de otras variables continuas y se da un valor puntual, pero se refieren a un intervalo.
Una variable aleatoria continua puede asumir todos los valores reales posibles en un intervalo $[a, b]$, siendo *a* y *b* tales que $-\infty < a < b < \infty$.

Resulta así que, aún en pequeños intervalos, *x* puede asumir infinitos valores no numerables y es prácticamente imposible obtener un valor significativo para la probabilidad puntual $P(x_i)$. En el campo continuo, entonces, la probabilidad se define para un intervalo, y no para un punto.

Se dice que una variable aleatoria *x* es del tipo continuo, si existe una función no negativa, $f(x)$, llamada *función de densidad* de *x*, que satisface la siguiente relación para todo valor real de *x*.

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

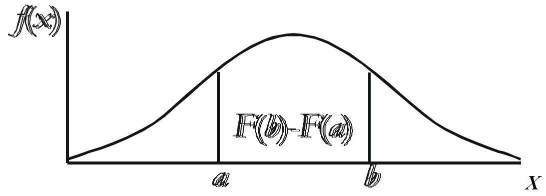
Siendo *F (x)* la *Función De Distribución* de *x*, que debe satisfacer a su vez, los requisitos de toda Función de Distribución.

De la expresión anterior, se deduce que $\dfrac{dF(x)}{dx} = f(x)$.

La función $f(x)$ es la derivada de *La Función De Distribución*; $f(x)$ no es una probabilidad, es la "*Densidad De La Probabilidad*", e indica la altura de la curva en el punto *x*. Cuando la variable aleatoria es continua la probabilidad se define para un intervalo, entonces:

$$P\ (a \leq x \leq b) = \int_a^b f(x)\,dx = F(b) - F(a)$$

Se convierte en un área bajo la curva.



$f(x_i)$, es la ordenada (la altura) de la curva $f(x)$ en el punto *x*, y su valor puede, ó no, ser 0. No obstante la Probabilidad para un punto específico *x*, será siempre igual a 0, cuando se trate de una variable continua.

$$P(x = x_i) = \int_{x_i}^{x_i} f(x)\,dx = F(x_i) - F(x_i) = 0$$

Pero esto no significa que el evento asociado a x, sea un evento imposible. De acuerdo a esto será indistinto trabajar con los signos $<$ ó $\leq$; $>$ ó $\geq$.

En síntesis:

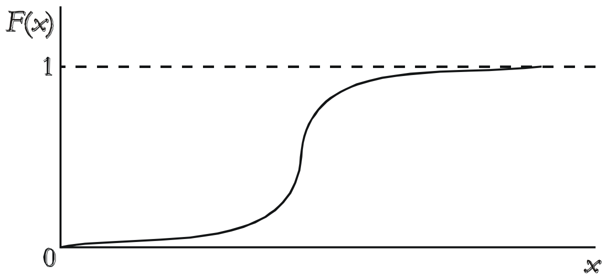
Una variable aleatoria es *continua* si su *Función De Distribución* $F(x)$, es continua y derivable, tal que su derivada primera dé la *Función De Densidad* $f(x)$ de la *Distribución De Probabilidades*. Conocida la *Función De Distribución* ó *De Acumulación* de una *Distribución De Probabilidades* se puede determinar la *Función De Densidad* con sólo derivar la primera y recíprocamente, conociendo la función de densidad se puede determinar la función de distribución, integrando aquella hasta un valor x de la variable aleatoria.

Finalmente la probabilidad correspondiente a un intervalo infinitesimal $(x,\ x + dx)$ se llama *Probabilidad Elemental y* se denota mediante

$$f(x)\,dx = d\,F(x)$$

En consecuencia $f(x)$ no es probabilidad de nada, pero $f(x)\,dx$, sí lo es.

Gráfica de una función de distribución de una variable aleatoria continua:



Una *Función De Distribución* Para Variables Aleatorias Continuas satisface las siguientes condiciones:

1) *F* (*x)* es una función monótona creciente (no decreciente), continua a la derecha de cada punto, pues es acumulación de probabilidades que son no negativas.

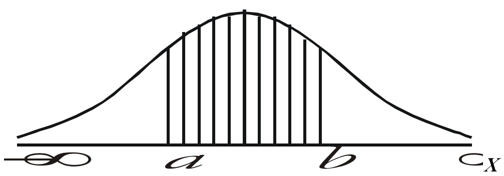
2) <u>Relaciones Asintóticas</u>

$F(-\infty) = Pr(x \leq -\infty) = 0$ *No hay valores p* (x_i) *para valores de* $x < -\infty$

$F(\infty) = Pr(x \leq \infty) = 1$ *Es la probabilidad del evento cierto*

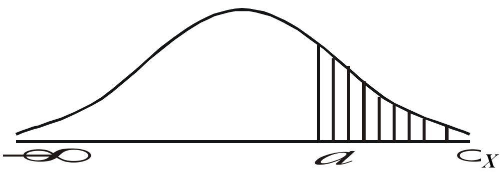
3) Dados dos números reales *a* y *b,* tales que $a < b$

$P(a \leq x \leq b) = F(b) - F(a)$



4) Dada una variable aleatoria *X* y un número real *a*

$P(x \geq a) = 1 - F(a)$



Para que *f(x)* sea función de densidad debe cumplir las siguientes condiciones básicas:

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

2º) $\int_{-\infty}^{x} f(x)\,dx = F(x)$, tal que $F(\infty) = 1$ y $F(-\infty) = 0$

<u>Ejemplo</u>:

Se supone una función $y=\frac{1}{2}x+2$, no negativa para valores de $x \geq -4$.

A partir de ella se deducirá una función de densidad, en cualquier intervalo entre $(-4;\infty)$, es decir en el intervalo positivo de dicha función, ya que la probabilidad es definida no negativa.

Interesa obtener la función de densidad en el intervalo (2, 6).

Para que la función propuesta sea función de densidad, deberá cumplir con las dos condiciones dadas anteriormente.

Veamos la primera:

1º) $\int_{-\infty}^{\infty} f(x)\,dx=1$, ó bien $\int_a^b f(x)\,dx=1$, si *a* y *b,* son los extremos del campo de variabilidad, ó intervalo dentro del cual se define la existencia de la variable aleatoria *X*. En el caso planteado, el intervalo (*a*, *b*) es (2, 6), por lo tanto, $\int_2^6 f(x)\,dx=1$ deberá cumplirse para que la *f* (x*)* sea función de densidad.

Entonces:

$$\int_2^6\left(\frac{x}{2}+2\right)dx=\int_2^6\frac{x}{2}\,dx+\int_2^6 2\,dx=\frac{1}{2}\int_2^6 x\,dx+2\int_2^6 dx=$$

$$=\frac{1}{2}\frac{x^2}{2}\Bigg|_2^6+2x\Bigg|_2^6=\frac{1}{2}\left(\frac{6^2}{2}-\frac{2^2}{2}\right)+2\ (6-2)=$$

$$=\frac{1}{2}\left(\frac{36}{2}-\frac{4}{2}\right)+8=\frac{1}{2}\frac{32}{2}+8=8+8=16$$

Es decir que $\int_2^6\left(\frac{x}{2}+2\right)dx=16$, por lo tanto se concluye que la función dada no es función de densidad.

Ahora bien, si se divide ambos miembros de la igualdad por 16, se tiene:

$$\frac{1}{16}\int_2^6\left(\frac{x}{2}+2\right)dx=1$$

Por lo tanto

$$\int_2^6\left(\frac{x}{32}+\frac{1}{8}\right)dx=1$$

Ahora sí, la función integrada $f(x)=\frac{x}{32}+\frac{1}{8}$ es función de densidad en el intervalo (2, 6) y se define en forma completa:

$$f(x)=\begin{cases}\dfrac{x}{32}+\dfrac{1}{8}; & 2\le x\le 6\\[2ex] 0 & \forall \text{ otro } x\end{cases}$$

Entonces, la primera condición queda demostrada, pasando a tratar la segunda de ellas:

$F(x)=\int_{-\infty}^{x} f(x)\,dx$, tal que $F(\infty) = 1$ y $F(-\infty) = 0$

El intervalo de integración será ahora (2, *x*), pues así fue definido el extremo inferior al considerar el intervalo (2, 6), entonces:

$$F(x)=\int_2^x f(x)dx=\int_2^x\left(\frac{x}{32}+\frac{1}{8}\right)dx=\frac{1}{32}\int_2^x x\,dx+\frac{1}{8}\int_2^x dx=$$

$$=\frac{1}{32}\frac{x^2}{2}\bigg|_2^x+\frac{1}{8}x\bigg|_2^x=\frac{1}{32}\left(\frac{x^2-4}{2}\right)+\frac{1}{8}(x-2)=\frac{x^2-4}{64}+\frac{x-2}{8}=$$

$$=\frac{x^2-4+8x-16}{64}=\frac{x^2}{64}+\frac{8x}{64}-\frac{20}{64}$$

Luego, $$F(x)=\frac{x^2}{64}+\frac{8x}{64}-\frac{20}{64}$$

Para verificar si la integral ha sido correctamente determinada la derivada de $F(x)$ debe dar la función integrada

$$f(x)=\frac{x}{32}+\frac{1}{8},\qquad \text{O sea:}\quad F'(x)=\frac{2x}{64}+\frac{8}{64}=\frac{x}{32}+\frac{1}{8}$$

Una vez determinada *F(x),* se verifica que:

a) $F(\infty)$ =1, o sea $F(6) = 1$, que se obtiene reemplazando en la función encontrada a *x* por 6.

O sea: $$F(6)=\frac{6^2}{64}+\frac{8(6)}{64}-\frac{20}{64}=\frac{36}{64}+\frac{48}{64}-\frac{20}{64}=\frac{64}{64}=1$$

b) $F(-\infty) = 0$, o sea $F(2) = 0$, que se obtiene reemplazando en la función encontrada a x por 2.

O sea: $$F(2)=\frac{2^2}{64}+\frac{8(2)}{64}-\frac{20}{64}=\frac{4}{64}+\frac{16}{64}-\frac{20}{64}=0$$

Comprobadas las propiedades, se puede definir la función de distribución en forma completa como:

$$F(x)=\begin{cases}\dfrac{x^2}{64}+\dfrac{8x}{64}-\dfrac{20}{64} & 2<x<6\\ 0 & x\le 2\\ 1 & x\ge 6\end{cases}$$

Gráficas

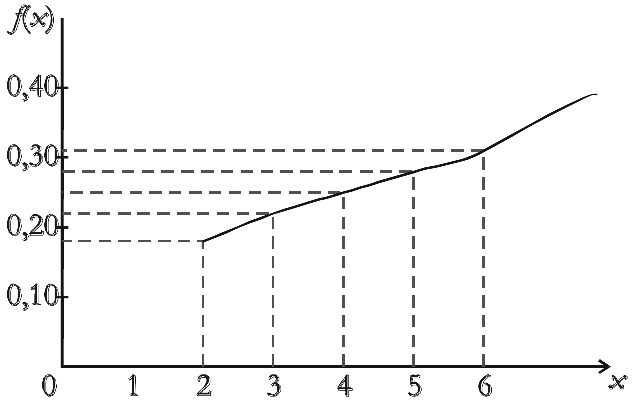
Hemos encontrado:

$$f(x)=\begin{cases}\dfrac{x}{32}+\dfrac{1}{8}; & 2\le x\le 6\\ 0 & \forall \text{ otro } x\end{cases} \quad \text{y} \quad F(x)=\begin{cases}\dfrac{x^2}{64}+\dfrac{8x}{64}-\dfrac{20}{64} & 2<x<6\\ 0 & x\le 2\\ 1 & x\ge 6\end{cases}$$

Para graficar damos a X valores desde 2 a 6 con lo que obtenemos los valores de la ordenada, en la abscisa marcamos los valores de X. Así,

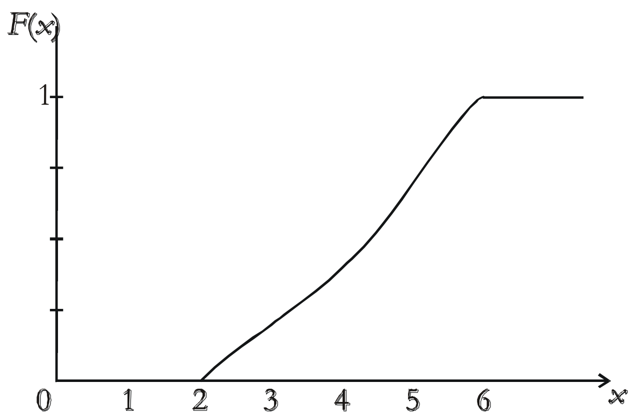
1) Función de densidad $f(x)$

x	$f(x) = x/32 + 1/8$
2	$\frac{2}{32}+\frac{1}{8}=\frac{2+4}{32}=\frac{6}{32}=0{,}1875$
3	$\frac{3}{32}+\frac{1}{8}=\frac{3+4}{32}=\frac{7}{32}=0{,}21875$
4	$\frac{4}{32}+\frac{1}{8}=\frac{4+4}{32}=\frac{8}{32}=0{,}25$
5	$\frac{5}{32}+\frac{1}{8}=\frac{5+4}{32}=\frac{9}{32}=0{,}28125$
6	$\frac{6}{32}+\frac{1}{8}=\frac{6+4}{32}=\frac{10}{32}=0{,}3125$



2) Función de Acumulación $F(x)$

x	$F(x)=\frac{x^2}{64}+\frac{8x}{64}-\frac{20}{64}$
2	$\frac{4}{64}+\frac{16}{64}-\frac{20}{64}=0$
3	$\frac{9}{64}+\frac{24}{64}-\frac{20}{64}=0{,}203125$
4	$\frac{16}{64}+\frac{32}{64}-\frac{20}{64}=0{,}4375$
5	$\frac{25}{64}+\frac{40}{64}-\frac{20}{64}=0{,}703125$
6	$\frac{36}{64}+\frac{48}{64}-\frac{20}{64}=1$



3. Los Parámetros en las Distribuciones de Probabilidad

Los parámetros, según lo visto, son medidas que se calculan con datos poblacionales en las distribuciones de frecuencias. En las distribuciones de probabilidad son calculados en base a los valores posibles que una variable puede asumir y se utilizan, conjuntamente con la función de probabilidad de esa variable, para caracterizar el fenómeno.

3.1. Esperanza Matemática

Así como para el caso de las distribuciones de frecuencias se hace referencia a la media aritmética, ahora, se considera para las distribuciones de probabilidad la esperanza matemática, valor esperado o simplemente media de una variable aleatoria. En las distribuciones de frecuencias la media aritmética es un valor que se ha verificado o realizado, es decir, se calcula luego de conocer los valores asumidos por la variable; en nuestro caso, la esperanza se puede determinar sin necesidad de practicar las observaciones, a condición que se conozca la distribución de probabilidades de la variable aleatoria y su valor es probable que se de, sin tener la seguridad o certeza que ello ha de ocurrir.

Es la suma de los productos de todos los posibles valores de la variable aleatoria por sus respectivas probabilidades. Cuando se trata de una variable aleatoria continua la suma se transforma en una integral.

Se representa por $E(\ )$, encerrando en los paréntesis a la variable aleatoria que se trate. Suele también ser representada por μ.

La distinción que se hace en la notación con respecto a la media aritmética se debe a que en este caso se trata de un parámetro de la distribución de probabilidad correspondiente a un espacio probabilístico y no a la media empírica de una distribución de frecuencias.

Se tiene entonces que:

Variables Aleatorias Discretas

$$E(x) = \mu_X = \sum_{i=1}^{N} x_i\, P_i \qquad \left|\ \begin{array}{l} i = 1,2,3,\ldots,N \quad \text{ó} \\ i = 1,2,3,\ldots \end{array}\right.$$

Variables Aleatorias Continuas

$$E(x) = \mu_X = \int_{\infty}^{\infty} x\, f(x)\, dx$$

Consideraciones:

1) N es ahora el total de valores posibles que la variable aleatoria puede asumir.

2) La Esperanza matemática es siempre un valor de la variable aleatoria x.

3) Está ubicada sobre el eje real de la x y es un parámetro de posición del colectivo o población, considerada en este caso, como el conjunto de valores posibles de la variable aleatoria.

Propiedades

1) Dada una variable aleatoria x, una función de ella $G(x)$, es también variable aleatoria.

<u>Ejemplo</u>:

En la tirada de un dado, la variable aleatoria puede asumir los valores x: 1, 2, 3, 4, 5, 6; con probabilidades iguales, $P_i = 1/6$.

Si se define

$G(x) = x^2$, los valores de esta función serán

$G(x) = x^2 = 1, 4, 9, 16, 25, 36$; y $P_i = 1/6$

x^2 que es una función de x, es también variable aleatoria, por lo tanto, su

Variables Aleatorias Discretas

$$E\left[G(x)\right] = \sum_{x} G(x)\, f(x)$$

Variables Aleatorias Continuas

y

$$E\left[G(x)\right] = \int_{-\infty}^{\infty} G(x)\, f(x)\, dx$$

2) La esperanza matemática de una constante, es la constante misma.

$$E(c) = c$$

Variables Aleatorias Discretas

$$E(c) = \sum_{i} c\, p_i = c \sum p_i = c.1 = c$$

Variables Aleatorias Continuas

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

3) La esperanza matemática de una constante por una variable aleatoria, es igual al producto de la constante por la esperanza matemática de la variable aleatoria.

$$E(c\,x) = c\,E(x)$$

Variables Aleatorias Discretas

$$E(c.x) = \sum_{i} c\, x_i\, p_i = c \sum x_i p_i = c\, E(x)$$

Variables Aleatorias Continuas

$$E(c.x) = \int_{-\infty}^{\infty} c\, x\, f(x)\, dx = c \int_{-\infty}^{\infty} x\, f(x)\, dx = c\, E(x)$$

4) La esperanza matemática de la suma algebraica de una variable aleatoria y una constante es igual a la esperanza de la variable aleatoria más/menos (suma algebraica) la constante.

$$E(x \pm c) = E(x) \pm c$$

Variables Aleatorias Discretas

$$E(x \pm c) = \sum_{i} (x_i \pm c)\, p_i = \sum x_i\, p_i \pm \sum c\, p_i = E(x) \pm c$$

Variables Aleatorias Continuas

$$E(x \pm c) = \int_{-\infty}^{\infty} (x \pm c)\, f(x)\, dx = \int_{-\infty}^{\infty} x\, f(x)\, dx \pm \int_{-\infty}^{\infty} c\, f(x)\, dx = E(x) \pm c$$

Corolario: De las propiedades anteriores, se deduce que: $E(a\,x \pm c) = a\,E(x) \pm c$

5) La esperanza matemática de la suma algebraica de variables aleatorias es igual a la suma algebraica de las esperanzas matemáticas de cada una de las variables.

$$E(x \pm x) = E(x) \pm E(x)$$

Esta propiedad puede extenderse a un número infinito de variables.

$$E\left[\sum_{i=1}^{N} x_{(i)}\right] = \sum_{i=1}^{N} E\left(x_i\right)$$

6) La esperanza matemática del producto de dos ó más variables aleatorias independientes, es igual al producto de las esperanzas matemáticas de dichas variables.

Si X e Y son variables aleatorias independientes, entonces:

$$E(x\ Y) = E(x)\ E(y)$$

Esta propiedad puede ser extendida al producto de un número finito de variables aleatorias

$$E\left(\prod_{i=1}^{N} x_{(i)}\right) = \prod_{i=1}^{N} E\left(x_{(i)}\right)$$

3.2. Varianza

La esperanza matemática es un parámetro que describe la tendencia central de una variable, la varianza permite conocer la concentración de los resultados alrededor de dicha esperanza.

Se simboliza y calcula:

Variables Aleatorias Discretas

$$V(x) = \sigma_X^2 = \sum_i [x_i - E(x)]^2 P_i = E\left(x^2\right) - [E(x)]^2 = \sum_i x_i^2 P_i - \left(\sum x_i P_i\right)$$

Variables Aleatorias Continuas

$$V(x) = \sigma_X^2 = \int_{\infty}^{\infty} [x_i - E(x)]^2\ f(x)\,dx = \int_{\infty}^{\infty} x^2\ f(x)\,dx - [E(x)]^2$$

Pudiendo reemplazar $E(x) = \mu$

<u>Propiedades:</u>

1) La Varianza de una variable aleatoria es no negativa.
2)

$$V(x) \geq 0$$

3) La Varianza de una constante es igual a 0.

$$V(c) = 0 \qquad \text{si } c = \text{constante}$$

$$V(c) = E\left\{[c - E(c)]^2\right\} = E(c - c)^2 = E(0) = 0$$

4) La Varianza de una constante por una variable, es igual al cuadrado de la constante por la varianza de la variable.

$$V(c\,x) = c^2\,V(x)$$

5) La Varianza de la suma algebraica de una variable y una constante es igual a la varianza de la variable.

$$V(x \pm c) = V(x)$$

6) La Varianza de una suma o diferencia de dos variable aleatoria independientes es igual a la suma de las varianzas de cada una de las variables.

$$V\left(x_{(1)} \pm x_{(2)}\right) = V\left(x_{(1)}\right) + V\left(x_{(2)}\right)$$

3.3. Desviación Estándar

Al igual que antes se define como la raíz cuadrada positiva de la varianza y su interpretación así como las razones para su uso son idénticas a las analizadas en el Unidad 3.

Se simboliza y calcula: $\sigma_x = \sqrt{\sigma_x^2}$

4. Momentos en las Distribuciones de Probabilidades.

Momento natural de orden k y Momento centrado de orden k

Además de la tendencia central y la variabilidad, de igual forma que una distribución de frecuencias, una función de probabilidad puede caracterizarse por el grado de asimetría y por el grado de agudeza. Medidas descriptivas de estas propiedades pueden definirse en términos de "momentos" de una distribución de probabilidades.

Momentos son simplemente la Expectativa o Esperanza de diferentes potencias de la variable aleatoria.

Los momentos se dividen en dos tipos; momentos naturales de orden k y momentos centrados de orden k, alrededor de la expectativa o esperanza.

Los primeros, tienen como objetivo calcular la esperanza de las variables aleatorias $x, x^2, \ldots, x^k$.

Así, el momento natural k-ésimo (alrededor del origen) de una variable aleatoria discreta x, se define como:

$$a_k = E(x^k) = \Sigma x_i^k\, p(x_i) \qquad i = 1,2,3,\ldots,N$$

Observamos que el primer momento natural k (alrededor del origen) es el valor esperado de una variable aleatoria.

Por otro lado, el momento natural k-ésimo (alrededor del origen) de una variable aleatoria continua x, se define como:

$$a_k = E(x^k) = \int x^k f(x)\,d(x) \qquad -\infty \leq x \leq \infty$$

Los momentos centrados, respecto de la Esperanza, cuando tienen orden par miden dispersión y cuando tienen orden impar se usan para medir asimetrías.

Así, para una variable aleatoria x, discreta, tenemos:

El momento centrado de orden k o momento de orden k con respecto de la media, es

$$m_k = \mu^k = [E(x_i - E(x)]^k \qquad i = 1,2,3,\ldots,N$$

A la vista de esta fórmula se ve fácilmente porque los momentos centrales de orden par miden dispersión y los de orden impar asimetrías: si el orden es par todas las diferencias con respecto de la media se elevan a una potencia par, haciéndose positivas, de modo que las distancias a izquierda y derecha de la media se suman. Por el contrario, cuando el exponente es impar las diferencias a izquierda y derecha se van cancelando, llevando a asimetrías positivas o negativas según el signo.

$$m_2 = \mu^2 = E[(x_i - \mu)^2] ,$$

Es la varianza de x. De igual forma que la media y la varianza miden la localización y dispersión de una distribución, respectivamente, así también momentos más altos miden otras propiedades de una distribución. Así,

$$m_3 = \mu^3 = E[(x_i - \mu)^3] ,$$

El tercer momento alrededor de la media, puede usarse para determinar si una distribución es simétrica o asimétrica. Puesto que todas las desviaciones en μ^3 son elevadas al cubo, las desviaciones negativas y positivas tenderán a anularse entre sí, dando

$\mu^3 = 0$, *si la distribución es simétrica alrededor de* μ .

$\mu^3 > 0$, *si la distribución es asimétrica a la derecha.*

$\mu^3 < 0$, *si la distribución es asimétrica a la izquierda.*

Al estar influido el tamaño de μ^3 por las unidades usadas para medir los valores de *x*, lo transforma en una deficiente medida de asimetría, es por ello que podemos utilizar una medida relativa, eliminando la dimensión de *x:*

$$K^3 = \frac{\mu_3}{\sigma^3}$$

De acuerdo al resultado que arroje este cociente, serán las conclusiones, así:

$K^3 = 0$; *indica simetría*

$K^3 > 0$; *indica asimetría positiva*
$K^3 < 0$; *indica asimetría negativa*

El cuarto momento alrededor de la expectativa es:

$$m_4 = \mu^4 = E[(x_i - \mu)^4]$$

Es siempre no negativo y puede usarse para mostrar el grado de agudeza o puntiagudez. También aquí podemos introducir una medida relativa, a los fines de eliminar la dimensionalidad, tal como:

$$K^4 = \frac{\mu_4}{\sigma^4} - 3$$

Esta medida nos representa la curtosis, agudeza, o puntiagudez. Así, con referencia a una distribución normal:

$K^4 = 0$, *la distribución es Mesocúrtica.*
$K^4 < 0$, *la distribución es Platicúrtica.*
$K^4 > 0$, *la distribución es Leptocúrtica.*

Por último, si x es una variable aleatoria continua, los momentos centrados se calculan de la siguiente manera:

$$m_k = \int [x - E(x)]^k f(x)\, d(x) \qquad -\infty \leq x \leq \infty$$

Diferencias entre Distribuciones de Probabilidad y Distribuciones de Frecuencias

Variables aleatorias ⇒ Distribuciones De Probabilidad					
Muestra el comportamiento de los valores posibles de una variable aleatoria					
Valores Posibles					
Variables Aleatorias Discretas			***Variables Aleatorias Continuas***		
y_i	$P(y_i)=$ $P(y=y_i)$	$P(y\leq y_i)$	$y'_{i-1}-y_i'$	$f(x)=\dfrac{dF(x)}{dx}$	$F(x)=\int_{-\infty}^{x} f(x)dx$
Valores Posibles	Función de Cuantía "Probabilidad Puntual"	Función de Acumulación "Probabilidad Acumulada"	Valores Posibles	Función de Densidad "Densidad de Probabilidad"	Función de Acumulación "Probabilidad Acumulada"

Variables ⇒ Distribuciones De Frecuencias									
Muestra el comportamiento de los valores observados de una variable en la *Población* o en la *Muestra*									
Valores Observados									
Variables Discretas					***Variables Continuas***				
y_i	n_i	h_i	N_i	H_i	$y'_{i-1}-y_i'$	n_i	h_i	N_i	H_i
Valores Observados	Frecuencias Puntuales		Frecuencias Acumuladas		Valores Observados	Frecuencias del Intervalo		Frecuencias Acumuladas Menores que o Mayores que	

Parámetros

Variables Aleatorias Discretas

$$E(y)=\sum_{i=1}^{N} y_i \; P(y_i)$$

$$\sigma_y^2=\sum_{i=1}^{N} y_i^{\,2} \; P(y_i)-[E(y)]^2$$

Variables Aleatorias Continuas

$$\int_{-\infty}^{\infty} y \;\; f(y)\, dy$$

$$\sigma_y^2=\int_{-\infty}^{\infty} y_i^{\,2} \;\; f(y_i)\, dy-[E(y)]^2$$

Parámetros (Población) – Estadígrafos (Muestra)

Variables Discretas o Continuas

$$M(y)=\sum_{i=1}^{m}\frac{y_i.n_i}{n}=\sum_{i=1}^{m} y_i.h_i$$

$$V(y)=\sum_{i=1}^{m}\frac{y_i^{\,2}.n_i}{n}-[M(y)]^2$$

CAPITULO Nº 6: Modelos Especiales de Probabilidad Variables Aleatorias Discretas

Objetivos Específicos

Que el estudiante:

Conozca *alguno de los diferentes modelos de probabilidad de variables aleatorias discretas.*

Distinga *las características de cada modelo, a lo fines de aplicarlos a situaciones reales.*

Determine *los valores posibles que pueden asumir las variables aleatorias involucradas en cada modelo.*

Reconozca *las funciones de probabilidad y de acumulación para el cálculo de probabilidades puntuales y acumuladas.*

Calcule *los parámetros en cada una de los modelos vistos.*

Analice *la simetría de las distribuciones tratadas.*

Conozca *el manejo de las tablas usuales y como las mismas fueron calculadas a través de sus respectivas funciones de probabilidad y de acumulación.*

Contenidos

1. *MODELOS ESPECIALES DE PROBABILIDAD VARIABLES ALEATORIAS DISCRETAS*

El objeto de una teoría, o un modelo, es explicar fenómenos y conductas.

Un modelo determinista nos permite decir que, dadas ciertas condiciones iniciales, es seguro que se obtendrán ciertos estados o resultados. En otras palabras, un modelo determinista es una explicación de causa y efecto: ciertas condiciones causan el estado subsiguiente.

En muchas situaciones no puede establecerse una clara relación, debido a la incertidumbre; aquí sólo podemos tener modelos probabilistas.

Un modelo probabilista, nos permite decir sólo que, dadas ciertas condiciones iniciales, ocurrirán ciertos estados con tales y tales probabilidades. Es decir, dadas las condiciones iniciales, un modelo probabilista nos permite deducir una distribución de probabilidades de posibles estados subsiguientes, que son valores de una variable aleatoria.

Entonces los Modelos Especiales De Probabilidad (Distribuciones De Probabilidades Para Variables Aleatorias) son modelos matemáticos apropiados para situaciones del mundo real en condiciones específicas. Son importantes porque ayudan a predecir la conducta de futuras repeticiones de un experimento, en otras palabras, un modelo probabilístico es una expresión matemática, resultante de un cúmulo de supuestos que tiene como propósito principal analizar los resultados de un determinado experimento aleatorio con el fin de poder predecir futuros resultados si dicho experimento se repite muchas veces.

Ejemplos:

Extraer al azar un número n de objetos, de una población en que hay solo dos características, por ejemplo hombres y mujeres. El modelo *Binomial* describe las probabilidades de obtener $0, 1, 2,.., n$ con una de las características.

El número de días de licencias médicas que se producen en una institución, o de fallas de un sistema computacional, en un mes, es aleatorio y podría representarse mediante un modelo *Poisson.*

La probabilidad de encontrar un ítem defectuoso al inspeccionar un número determinado de ítems de un lote pequeño, puede representarse mediante un modelo *Hipergeométrico.*

Como puede observarse en los casos planteados hay diferentes modelos, todos ellos, aplicables en diferentes situaciones, según distinguimos seguidamente:

Tipo de Variable	*Tamaño de la Población*	*Tipo de Muestreo*	*Tamaño de la Muestra*	*Comportamiento de la Probabilidad*	*Estadísticamente*	*Modelo a seleccionar*
Discretas	*Finita*	*MCR*	*Cualquiera*	*Constante*	*Independencia*	*Binomial*
		MSR	$n > 5\%$ *s/N*	*Variable*	*Dependencia*	*Hipergeométrica*
	Infinita	*MCR*	*Cualquiera*	*Constante*	*Independencia*	*Binomial*
		MSR	$n \leq 5\%$ *s/N*	*Constante*	*Independencia*	*Binomial*
	A partir de $x \approx B(n,P)$ *si n es grande y P es pequeña*					*Poisson*
	Modelo Uniforme Discreto					

La aplicación del modelo adecuado nos permitirá determinar el total de valores posibles para la variable aleatoria y sus correspondientes probabilidades de ocurrencia a través de sus Funciones De Probabilidad y Funciones De Distribución.

Las variables aleatorias discretas, pueden tomar valores que difieren entre sí en valores finitos.

Supondremos a lo largo de todo el desarrollo de la presente Unidad, que los datos son resultado de un muestreo aleatorio simple, es decir toda posible muestra de *n* unidades tiene la misma probabilidad de ser elegida, o bien todo elemento de la población tiene idéntica probabilidad de ser seleccionado para conformar la muestra.

Los Modelos Especiales De Probabilidad para Variables Discretas se basan en el Proceso Aleatorio De Bernoulli, o Distribución De Bipuntual.

1.1. <u>MODELO DE BERNOULLI</u>

Características

Hay un grupo considerable de fenómenos aleatorios que tienen la distribución de probabilidad más sencilla, pues sólo hay dos sucesos ó resultados, mutuamente exclusivos y colectivamente exhaustivos. Es decir se aplica a una variable que puede asumir sólo dos valores, por ello se habla de población dicotómica.

<u>Ejemplo</u>:

El 10% de los trabajadores del país está desempleado, ¿Cuál es la probabilidad de seleccionar un individuo al azar y esté desempleado?

Analizando el ejemplo planteado, se observa que sólo existen dos posibles resultados, Desempleado (característica estudiada) o Empleado (característica no estudiada). Luego, generamos la variable aleatoria, dando el valor 1 a la característica estudiada con probabilidad P, y le damos el valor 0 a la característica no estudiada, con probabilidad Q =1-P.

Característica	*Probabilidad*	*Variable Aleatoria* y_i	*Probabilidad* $P(y=y_i)$
Desempleado	*0,10*	*1*	$P(y=1)=P=0{,}10$
Empleado	*0,90*	*0*	$P(y=0)=Q=1-P=0{,}90$

Para normalizar la terminología que describe estos y otros muchos procesos semejantes, diremos que uno de los dos resultados posibles es un éxito, que representa a la característica estudiada y el otro un fallo, que representa a la característica no estudiada.

Los dos sucesos, éxito y fallo, son de naturaleza cualitativa, pero se pueden convertir en cuantitativos asignándoles el valor 1 al éxito y el valor 0 al fallo.

Sea y una variable aleatoria de Bernoulli con los posibles valores 0 y 1. Sea así mismo P, la probabilidad de éxito y $Q = 1 - P$ la probabilidad de fallo.

Entonces el Modelo Probabilístico De Bernoulli, Proceso de Bernoulli o Distribución Bipuntual tendrá la forma general:

y	$P(y)$
0	$Q = 1 - P$
1	P

Función de Probabilidad: Función de Cuantía

$$P(y) = \begin{vmatrix} P^y (1-P)^{1-y} & \quad y = 0,1 \\ 0 & \quad \forall \text{ otro } y \end{vmatrix}$$

Como tiene solamente dos clases de sucesos elementales, también se la conoce como modelo probabilístico de dos puntos.

Parámetros

Igual que otras variables aleatorias, la variable aleatoria de Bernoulli, tiene su Esperanza, su Varianza y su Desviación Estándar.

Así:

$$E(y) = \sum_{Y=0}^{1} y\, P(y) = 0\,Q + 1\,P = \underline{P}$$

$$V(y) = \sum_{Y=0}^{1} y^2\, P(y) - [E(y)]^2 = \left[0^2\,Q + 1^2\,P\right] - P^2 = P - P^2 = \underline{PQ}$$

$$\sigma_\gamma = \sqrt{PQ}$$

Es decir, la media de una variable aleatoria Bernoulli es igual a la probabilidad de un éxito, y la varianza es igual a la probabilidad de un éxito multiplicada por la de un fallo.

La media "P", es un parámetro de población, llamado la Proporción Poblacional, y es la razón esperada del número de éxitos y el tamaño de la población.

$$\frac{X}{N} = \frac{\text{Número de Éxitos en la Población}}{\text{Tamaño de la Población}}$$

Podemos saber en todo momento si un proceso es Bernoulli, pero puede que no conozcamos el valor del parámetro *P*. Con un dado perfecto podemos saber si el proceso es Bernoulli con *P* = 1/6 de éxitos (ejemplo as) y *Q* = 5/6 de fracasos (otros números).

Pero si se nos da un dado, y se nos dice que no es perfecto, el proceso (echar el dado) puede ser aún Bernoulli, pero el parámetro *P* es ahora desconocido por nosotros.

Así, podemos tener un proceso Bernoulli con un parámetro conocido ó uno desconocido.

Calculamos para el ejemplo dado, la Esperanza matemática y la Varianza y la Desviación Estándar

$$E(y) = P = 0,90$$

$$\sigma_y^2 = PQ = 0,90x0,10 = 0,09$$

$$\sigma_y = \sqrt{PQ} = \sqrt{0,90 \times 0,10} = \sqrt{0,09} = 0,3$$

1.2. MODELO BINOMIAL: Número De Éxitos En n Pruebas

Características

Supóngase que una muestra aleatoria de n observaciones se toma por el Proceso de Bernoulli. Sea esta muestra: $y_1, y_2, \ldots, y_n$

Como *y* toma el valor 1 si el resultado de una prueba es un éxito y 0 si es un fallo, el número de unos en *n* pruebas es simplemente $y_1 + y_2 + \ldots + y_n$ o sea:

$$\sum_{i=1}^{n} y_i$$

Se introduce ahora una nueva variable aleatoria *x* que se llama *Variable Aleatoria Binomial* y que se utiliza para designar el número de éxitos en una muestra de n observaciones tomadas por el Proceso de Bernoulli, esto es:

$$x = \sum_{i=1}^{n} y_i$$

Esta distribución se basa, según vimos anteriormente en el supuesto de:

- Población Infinita ó Finita, Muestreo con Reemplazo.

- Población Infinita, Muestreo sin Reemplazo con n ≤ 5 % de N.

De lo dicho podemos inferir:

1. Se repite una prueba simple un número *n* de veces, cada una de ellas independientes entre sí y bajo las mismas condiciones.

2. En cada prueba sólo pueden presentarse dos alternativas mutuamente excluyentes: "Éxito y Fracaso", con probabilidades asociadas P y $Q = 1 - P$, respectivamente, siendo $P + Q = 1$, es decir que cada prueba tiene distribución bipuntual.
3. Por ser cada prueba independiente de las demás y realizarse en las mismas condiciones, las probabilidades P y Q permanecen constantes a lo largo de todo el experimento.
4. En cada prueba se centra la atención en determinar si se presenta un éxito o un fracaso con el fin de totalizarlos al final de las n pruebas.

Entonces, un determinado experimento que presenta las características anteriores tiene Distribución Binomial o responde al Modelo Binomial y la variable Binomial cuenta con n +1 valores posibles, si el número de pruebas repetidas es n, donde el mínimo valor es 0 y el máximo valor es n.

Si a partir de un experimento aleatorio que tiene sólo dos resultados posibles, mutuamente excluyentes y colectivamente exhaustivos, "éxito" y "fracaso", donde P es la probabilidad de obtener éxito en cada repetición, se realizan n repeticiones independientes, la Distribución del Número de Éxitos, x, resultante se denomina Distribución Binomial.

$$x \approx B\,(n,\, P)$$

La Distribución Binomial está completamente definida por n y por P, los parámetros de la distribución.

El valor de n y P especifican una distribución particular entre todas las posibles Distribuciones Binomiales. Si x es una variable aleatoria Binomial con n y P como parámetros de su distribución, se puede, entonces decir brevemente que "x tiene la distribución B(n, P)", o sencillamente que "x es B(n, P)".

Función de Probabilidad: Función de Cuantía

Por tratarse de una variable discreta la Función de Probabilidad se llama Función de Cuantía.

Nuestro estudio consiste en determinar una función que nos dé en forma directa las probabilidades respectivas para los valores posibles de la variable aleatoria Binomial, es decir de 0 hasta n, según mencionamos anteriormente.

<u>Ejemplo</u>:

¿Cuál es la probabilidad de obtener dos caras cuando se tira tres veces una moneda con probabilidad 1/2 de cara?

En general: ¿Cuál es la probabilidad de obtener x éxitos al realizar n pruebas repetidas con probabilidad constante P de presentación del éxito?, que simbolizaremos por:

$$P(x = x_i\,;\, n\,;\, P)$$

Y el problema reside en determinar la estructura analítica de esta función, donde n y P son los parámetros que la caracterizan, es decir especifican una distribución particular entre todas las posibles distribuciones binomiales.

A este problema lo resolvemos con la Función de Cuantía de la Distribución Binomial.

Entonces, la probabilidad de que se presenten x éxitos en n pruebas repetidas e independientes, con probabilidad constante P para los éxitos es:

$$\Pr(x = x_i, n, P) = \left| \begin{array}{ll} C_n^x \; P^x \; Q^{n-x} & x = 0,1,2,\ldots,n \\ 0 & \forall \text{ otro valor de } x \end{array} \right.$$

Donde $$C_n^x = \frac{n!}{x!(n-x)!}$$

Veamos a través de un ejemplo sencillo como llegamos a la anterior expresión para la función de cuantía:

Supongamos que se juegan tres monedas, todas perfectas y que interesa la cantidad de caras que se presentan. Del enunciado anterior observamos:

1- $n = 3$
2- $P = Q = 0{,}50$, pues suponemos monedas perfectas
3- Los resultados son independientes dado que P = 0,50 permanece constante a lo largo de todo el experimento.

Entonces, la variable aleatoria x, que designa el número de caras obtenidas, asumirá $n + 1$ valores posibles, que para establecerlos trabajamos como sigue:

Al tirar las tres monedas pueden producirse los siguientes resultados posibles (eventos elementales) que constituyen el espacio probabilístico:

Ω = [(sss); (css); (scs); (ssc); (ccs); (csc); (scc); (ccc)]

Cada evento elemental dará lugar a un valor para la variable aleatoria x, que insistimos representa el número de caras (éxitos).

Así en (sss) no se presenta cara, por lo tanto x asume el valor cero; en (css), se presenta una cara, por lo tanto x asume el valor 1 y así sucesivamente.

Entonces, resumiendo tenemos:

Evento elemental		Nº de caras (*X*)
css	000	0
css	100	1
scs	010	1
ssc	001	1
ccs	110	2
csc	101	2
scc	011	2
ccc	111	3
	- Tabla 1 -	

De manera tal que los distintos valores que x puede adoptar son:

0, 1, 2 y 3 ($n+1 = 3 + 1$).

Nos interesa encontrar la probabilidad de cada uno de estos cuatro valores posibles.

Para facilitar el desarrollo llamemos y_1, y_2, y y_3 a los tres resultados. Por otro lado recordemos que al éxito lo representábamos con el número 1 (con probabilidad P asociada) y al fallo con el número 0 (con probabilidad Q asociada), por lo cual los eventos elementales quedarán indicados con 0 y 1 según sea fallo (sello) o éxito (cara) respectivamente.

Por consiguiente, se tiene:

$$
\begin{aligned}
P(x=0) &= P(y_1=0 \text{ y } y_2=0 \text{ y } y_3=0)\\
&= P(y_1=0)\times P(y_2=0)\times P(y_3=0), \quad \textit{pues los resultados son independientes}\\
&= Q\times Q\times Q = Q^3 = \left(\tfrac{1}{2}\right)^3 = \frac{1}{8}
\end{aligned}
$$

$$
\begin{aligned}
P(x=1) &= P(y_1=1 \text{ y } y_2=0 \text{ y } y_3=0) \text{ ó } P(y_1=0 \text{ y } y_2=1 \text{ y } y_3=0) \text{ ó }\\
&\quad P(y_1=0 \text{ y } y_2=0 \text{ y } y_3=1)\\
&= [P(y_1=1)\times P(y_2=0)\times P(y_3=0)] + [P(y_1=0)\times P(y_2=1)\times P(y_3=0)] +\\
&\quad [P(y_1=0)\times P(y_2=0)\times P(y_3=1)]\\
&= (P\times Q\times Q) + (Q\times P\times Q) + (Q\times Q\times P) = PQ^2 + PQ^2 + PQ^2 = \underline{3PQ^2}\\
&= 3\,\frac{1}{2}\,\left(\frac{1}{2}\right)^{2} = 3\left(\frac{1}{2}\right)^3 = \frac{3}{8}
\end{aligned}
$$

$$
\begin{aligned}
P(x=2) &= P(y_1=1 \text{ y } y_2=1 \text{ y } y_3=0) \text{ ó } P(y_1=1 \text{ y } y_2=0 \text{ y } y_3=1) \text{ ó }\\
&\quad P(y_1=0 \text{ y } y_2=1 \text{ y } y_3=1)\\
&= [P(y_1=1)\times P(y_2=1)\times P(y_3=0)] + [P(y_1=1)\times P(y_2=0)\times P(y_3=1)] +\\
&\quad [P(y_1=0)\times P(y_2=1)\times P(y_3=1)]\\
&= (P\times P\times Q) + (P\times Q\times P) + (Q\times P\times P) = P^2Q + P^2Q + P^2Q = \underline{3P^2Q}\\
&= 3\left(\frac{1}{2}\right)^2 \frac{1}{2} = 3\left(\frac{1}{2}\right)^3 = \frac{3}{8}
\end{aligned}
$$

$$
\begin{aligned}
P(x=3) &= P(y_1 = 1 \text{ y } y_2 = 1 \text{ y } y_3 = 1) \\
&= P(y_1 = 1) \times P(y_2 = 1) \times P(y_3 = 1) \\
&= P \times P \times P = P^3 = \left(\frac{1}{2}\right)^3 = \frac{1}{8}
\end{aligned}
$$

Obsérvese que $P(x = 0)$ es la probabilidad de obtener tres sellos, ó Q^3, ninguna cara, o sea P^0. Como $P^0 = 1$, $P(x = 0)$ se puede expresar por Q^3P^0. Análogamente, $P(x = 3)$ se puede expresar por P^3Q^0. Por otra parte, P^0Q^3 y $P^3\ Q^0$ no se alteran al multiplicarlas por un factor 1. Así:

$$P(x = 0) = 1\ P^0\ Q^3$$

$$P(x = 1) = 3\ P^1\ Q^2$$

$$P(x = 2) = 3\ P^2\ Q^1$$

$$P(x = 3) = 1\ P^3\ Q^0$$

Los factores 1, 3, 3 y 1 se llaman *coeficientes binomiales* y son combinaciones de *n* elementos tomándolos de a *x* cada vez.

En el ejemplo se tiene:

$$C_3^0 = \frac{3!}{0!\,3!} = 1 \qquad C_3^1 = \frac{3!}{1!\,2!} = 3 \qquad C_3^2 = \frac{3!}{2!\,1!} = 3 \qquad C_3^3 = \frac{3!}{3!\,0!} = 1$$

En base a esto, las expresiones dadas en (1) pueden generalizarse así:

$$
\begin{aligned}
&P(x=0) = C_n^0 P^0 Q^{n-0} \\
&P(x=1) = C_n^1 P^1 Q^{n-1} \\
&P(x=2) = C_n^2 P^2 Q^{n-2} \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&P(x=n) = C_n^n P^n Q^{n-n}
\end{aligned}
$$

Así, sustituyendo 0, 1, 2,..., *n* por *x*, se obtiene:

$$P(x = x_i) = C_n^x P^x Q^{n-x}$$

Que es la expresión dada al inicio del tema.

Entonces, la función de Cuantía para el número de caras (*x*) en el juego de tres monedas ideales, es:

$x = x_i$	$P(x = x_i \quad n = 3 \quad P = 1/2)$
0	1/8
1	3/8
2	3/8
3	1/8

Tal vez no resulte claro que cada una de las ocho sucesiones de la Tabla 1 sea una combinación en vez de una permutación. Tomemos css por ejemplo. Aquí es y_1 = 1, y_2 = 0 y y_3 = 0. Si tuviera importancia el orden de ocurrencia de estas tres observaciones, entonces otras disposiciones ordenadas de estas tres observaciones, tales como (y_2= 0, y_1 = 1, y_3= 0), (y_3= 0, y_2 = 0, y_1= 1), (y_2= 0, y_3 = 0, y_3= 1) y (y_3= 0, y_1 = 0, y_3= 0), se considerarían diferentes unas de otras. Pero en realidad son idénticas puesto que no interesa el orden en que están dispuestas, lo que hace que constituyan una combinación.

Función de Distribución o Función de Acumulación

Para una variable discreta hemos definido la Función De Distribución como

$$F(x) = P(x \leq x_i) = \sum_{x=0}^{x_i} P(x)$$

En nuestro caso será:

$$F(x) = P(x \leq x_i, n, P) = \sum_{x=0}^{x_i} C_n^x P^x Q^{n-x}$$

Por lo tanto, la probabilidad de que la variable aleatoria tome un valor cualquiera en el intervalo (*a*, *b*) para *a* < *b*, será:

$$P(a \leq x \leq b) = \sum_{x=a}^{b} C_n^x P^x Q^{n-x}$$

De acuerdo a las propiedades de esta función, tendremos:

$$F(\infty) = F(n) = \sum_{x=0}^{n} C_n^x P^x Q^{n-x} = 1$$

Ejemplo:

Una compañía recibe un gran cargamento de artículos, y decide aceptar el envío si en una muestra aleatoria de veinte artículos no hay más de uno defectuoso. Es decir, se acepta el cargamento si el número de artículos defectuosos es cero o uno, por lo que si Pr (x) es la función de probabilidad del número x de artículos defectuosos en la muestra, tenemos:

Pr (aceptar el cargamento) = Pr (x=0) + Pr (x=1)

Supongamos que la proporción de artículos defectuosos en el cargamento es P = 0,1.

Para n= 20, en la Tabla 1, encontramos que las probabilidades de cero y un artículos defectuosos en la muestra son, respectivamente, Pr (x=0 n=20 P=0,1) = 0,1216 y

Pr (x=1 n=20 P=0,1) = 0,2702. Por tanto, con esta regla de decisión, la probabilidad de que la compañía acepte el envío es

P (aceptar el cargamento) = 0,1216 + 0,2702 = 0,3918

Análogamente, si el 20% de los artículos del cargamento son defectuosos, es decir, si P=0,2, entonces,

P (aceptar el cargamento) = 0,0115 + 0,0576 = 0,0691

Y para P= 0,3

P (aceptar el cargamento) = 0,0008 + 0,0068 = 0,0076

Parámetros

Esperanza

La variable aleatoria *x*, con distribución Binomial es la suma de *n* variables "*y*", con Distribución Bipuntual, o sea que cada y_i tiene $E(y) = P$ y $V(y) = PQ$ por lo tanto:

Si $x = \sum_{i=1}^{n} y_i \Rightarrow E(x) = E\left(\sum_{i=1}^{n} y_i\right)$ y por propiedad de esperanza matemática:

$E(x) = \sum_{i=1}^{n} E(y)$, y siendo $E(y) = P$, nos queda: $E(x) = \sum_{i=1}^{n} P$, pero la suma de una constante es *n* veces la constante $\therefore$ $\boxed{E(x) = nP}$

Varianza y Desviación Estándar

$V(x) = V\left[\sum_{i=1}^{n} y_i\right]$ Y por ser las y_i, independientes

$$V(x) = \sum_{i=1}^{n} V(y)$$

$$V(x) = \sum_{i=1}^{n} PQ$$

$$\boxed{V(x) = nPQ} \Rightarrow \boxed{\sigma_x = \sqrt{nPQ}}$$

Configuración

La forma de la Distribución Binomial, depende de los valores de P y *n*.

1- Si P = Q = 0,50, la distribución será simétrica, independientemente del valor de n.

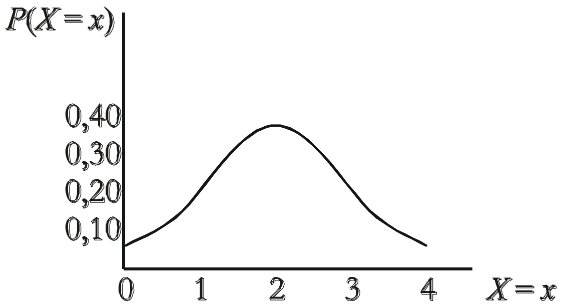
Ejemplo:

$n = 4$ y $P = Q = 0{,}50$

Entonces, la Distribución De Probabilidad será:

$x = x_i$	$P(x) = P(x = x_i \quad n = 4 \; P = 0{,}50) = C_4^x (0{,}50)^x (0{,}50)^{4-x}$
0	0,0625
1	0,2500
2	0,3750
3	0,2500
4	0,0625

Gráficamente:

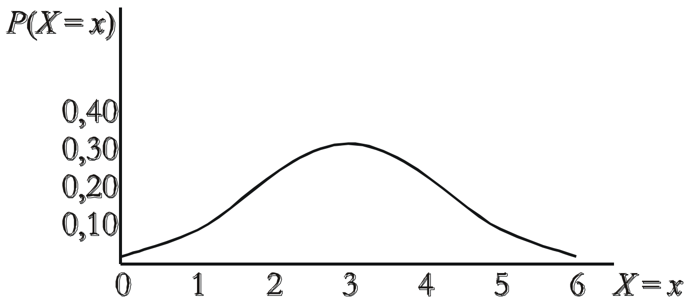


$n = 6$ y $P = Q = 0{,}50$

Entonces, la Distribución De Probabilidad será:

$x = x_i$	$P(x) = P(x = x_i \quad n = 6 \quad P = 0{,}50) = C_6^x (0{,}50)^x (0{,}50)^{6-x}$
0	0,0156
1	0,0938
2	0,2344
3	0,3125
4	0,2344
5	0,0938
6	0,0156

Gráficamente:



2. Si $P \neq Q$, la distribución será asimétrica.

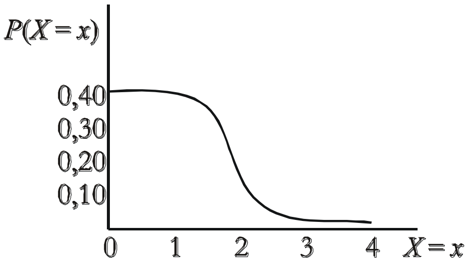
Si $P < Q$, la distribución será asimétrica hacia la derecha a de asimetría positiva.

Ejemplo:

Consideremos $n = 4$ y $P = 0{,}20$, entonces la Distribución De Probabilidad será:

$X = x_i$	$P(x) = P(x = x_i \quad n = 4 \quad P = 0{,}20) = C_4^x (0{,}20)^x (0{,}80)^{4-x}$
0	0,4096
1	0,4096
2	0,1536
3	0,0256
4	0,0016

Gráficamente:



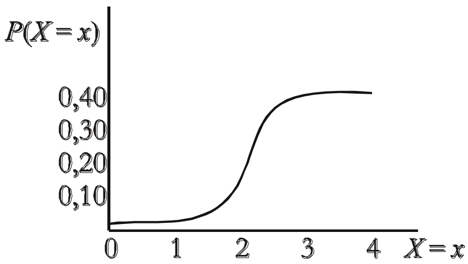
Si P > Q, será asimétrica hacia la izquierda, o de asimétrica negativa.

Ejemplo:

Consideremos $n = 4$ y $P = 0{,}80$, entonces la Distribución De Probabilidad será:

$X = x_i$	$P(x) = P(x = x_i \quad n = 4 \quad P = 0{,}80) = C_4^x (0{,}80)^x (0{,}20)^{4-x}$
0	0,0016
1	0,0256
2	0,1536
3	0,4096
4	0,4096

Gráficamente:



Sin embargo, al incrementar el tamaño de la muestra (n), la distribución se hace menos asimétrica. Al tender n a infinito ($n \to \infty$) se aproximará a la simetría independiente de la diferencia entre P y Q.

Tablas Usuales

La distribución, es una distribución muestral, ya que la variable *x* es un estadígrafo muestral, que, como la media muestral, se obtiene de una muestra de *n* observaciones.

Pero cuando el valor de *n* es grande, el cálculo manual de $C_n^x P^x Q^{n-x}$ se hace complicado. Es por ello, y ante la carencia de tecnología individual en el aula, es que para la resolución de ejercicios, utilizaremos las tablas de distribuciones binomiales para varios valores de *n* y *P*.

Existen distintos formatos de tablas que están de acuerdo con lo que cada autor considera mas conveniente, y generalmente cada texto cuenta con un apéndice en el cual se presentan un conjunto de tablas para distintas distribuciones.

Las que utilizaremos, presentan para la Distribución Binomial dos tablas:

Tabla I, referida a la *Función de Probabilidad*, o sea $Pr(x = x_i, n, P)$. Las entradas en la tabla son valores de $C_n^x P^x Q^{n-x}$ para los valores indicados de *n*, *x* y *P*.

Ejemplo:

$$\Pr(x = 3 \quad n = 5 \quad P = 0{,}20) = \underline{0{,}0512}$$

Para resolver esta probabilidad entramos en la Tabla I, buscando $n = 5$. Entre este grupo de valores ubicamos al $x = 3$ y en forma horizontal buscamos la intersección con $P = 0{,}20$.

En este punto, en el cuerpo de la tabla, aparece 0,0512 que constituye la respuesta al problema.

Observaciones:

a) Cuando $P <$ 0,50 no se encontrase en la tabla por ser un número comprendido entre los que figuran, el valor de $C_n^x P^x Q^{n-x}$ para un *n*, *x* y *P* dados, se obtiene por resolución de la fórmula.

Ejemplo:

$$\Pr(x = 2 \quad n = 4 \quad P = 0{,}23) = C_4^2 P^2 Q^{4-2} = \frac{4!}{2!\,2!}(0{,}23)^2(0{,}77)^2 =$$

$$= \frac{4 \times 3 \times 2 \times 1}{(2 \times 1)(2 \times 1)} \times 0{,}0529 \times 0{,}5929 = 0{,}1881864$$

b) Cuando $P >$ 0,50, (y complementario de los que figuran en tabla) el valor de $C_n^x P^x Q^{n-x}$ para un *n*, *x* y *P* dados, se obtiene encontrando el valor de la celda para el n dado, con $n - x$ en lugar del *x* dado y $1 - P$ en lugar de *P*.

Ejemplo:

$$\Pr(x = 2 \quad n = 6 \quad P = 0{,}80)$$

Al no encontrar *P* = 0,80 en tabla deberíamos resolver aplicando la fórmula, o sea:

$$\Pr(x = 2 \quad n = 6 \quad P = 0{,}80) = C_6^2 (0{,}80)^2 (0{,}20)^4 = 0{,}01536 \cong \underline{0{,}0154}$$

Pero a los fines de utilizar la tabla, buscamos:

$\Pr(x' = n - X = 4 \quad n = 6 \quad 1 - P = 0{,}20) = \underline{0{,}0154}$, o sea idéntico resultado.

c) Cuando $P > 0{,}50$ (y no complementario de los que figuran en tabla), se aplica el mismo criterio que en a).

Tabla II, referida a la *Función de Acumulación*, o sea *Pr(x ≤ xᵢ, n, P).* Es decir que acumula desde 0 hasta el valor dado de *x*, para *n* y *P* indicados.

Ejemplo:

1) $\Pr(x \geq 4 \quad n = 6 \quad P = 0{,}10)$

No podemos resolver esta probabilidad en forma directa, pues la tabla acumulada para valores ≥ *X*. Luego procederemos de la siguiente manera:

$$\begin{aligned}\Pr(x \geq 4 \quad n = 6 \quad P = 0{,}10) &= 1 - \Pr(X \leq 3 \quad n = 6 \quad P = 0{,}10) \\ &= 1 - 0{,}9987 \\ &= 0{,}0013\end{aligned}$$

2) $\Pr(x \leq 5 \quad n = 10 \quad P = 0{,}50) = 0{,}6230$

Ahora entramos directamente en tabla, buscando *n* = 10, luego en entre este grupo de valores el 5 y en forma horizontal la intersección con *P* = 0,50.

Observaciones:

a) Cuando $P < 0{,}50$ no se encontrase en tabla por ser un número comprendido entre los que figuran, el valor $P(X \leq x_i)$ para una *n,* y *P* dados se obtiene por resolución de la fórmula:

Ejemplo:

$$Pr(x \leq 1 \quad n = 4 \quad P = 0{,}23) = \sum_{x=0}^{1} C_4^x \, 0{,}23^x (0{,}77)^{4-x}$$

b) Cuando $P > 0{,}50$ (y complementario de los que figuran en la tabla) el valor $P(x_i \leq x_i)$ para un *n*, y *P* dados, se encontrara con *x'* en lugar del valor *x* dado y 1 – *P*, en lugar del valor *P* dado.

Ejemplo:

$$\Pr(x \leq 3 \quad n = 10 \quad P = 0{,}60) =$$

 $x;\ P=0{,}60$

 $x';\ 1-P=0{,}40$

$$\begin{aligned}\Pr(x'\geq 7 \quad n=10 \quad 1-P=0{,}40) &= 1-\Pr(x'\leq 6 \quad n=10 \quad 1-P=0{,}40)\\ &= 1-0{,}9452\\ &= 0{,}0548\end{aligned}$$

c) Cuando $P >$ 0,50 (y no complementario de los que figuran en tabla), se aplica el mismo criterio que en a).

<u>Ejercicios</u>:

1) Para: a) n = 3; b) n = 5, y siendo P = Q =1/2, hallar la función de probabilidad y la función de acumulación de x.

a) n = 3 y P = 0,50 = 1/2

$x=x_i$	$\Pr(x=x_i \quad n=3 \quad P=0{,}50)=C_3^x P^x Q^{3-x}$	$\Pr(x\leq x_i \quad n=3 \quad P=0{,}50)=\sum_{x=0}^{x_i} C_3^x P^x Q^{3-x}$
0	$\Pr(x=0 \quad n=3 \quad P=0{,}50)=0{,}1250=1/8$	$\Pr(X\leq 0)=\Pr(X=0)=0{,}1250=1/8$
1	$\Pr(x=1 \quad n=3 \quad P=0{,}50)=0{,}3750=3/8$	$\Pr(X\leq 1)=0{,}50=4/8$
2	$\Pr(x=2 \quad n=3 \quad P=0{,}50)=0{,}3750=3/8$	$\Pr(X\leq 2)=0{,}8750=7/8$
3	$\Pr(x=3 \quad n=3 \quad P=0{,}50)=0{,}1250=1/8$	$\Pr(X\leq 3)=1=8/8$
	$\sum_{x=0}^{3} P(x)=1$	

b) n = 5 y P = 0,50 = ½

$x=x_i$	$P(x=x_i)=C_5^x P^x Q^{5-x}$	$P(x\leq x_i)=\sum_{x=0}^{x_i} C_5^x P^x Q^{5-x}$
0	0,0313 = 1/32	0,0313
1	0,1563 = 5/32	0,1875
2	0,3125 = 10/32	0,50
3	0,3125 = 10/32	0,8125
4	0,1563 = 5/32	0,9687
5	0,0313 = 1/32	1
	$\sum_{x=0}^{5} P(x_i)=1$	

2) Existen 4 categorías mutuamente excluyentes de sangre humana: *Grupo* 0, *Grupo* A, *Grupo* B y *Grupo* AB. En cierta población las probabilidades de estos grupos son, respectivamente: 0,46; 0,40; 0,11; y 0,03. Suponiendo una población muy grande, hallar la probabilidad de que en una muestra tomada al azar, de 4 individuos, se encuentren:

a) Dos personas del Grupo 0

b) Ninguna persona del Grupo AB

c) Definir la variable aleatoria como el número de personas del Grupo A, que puede tener la muestra y determinar su función de probabilidad.
d) Calcular la Esperanza y la Varianza de la Variable, del punto c).

La distribución que corresponde aplicar es la Distribución Binomial, dado que hay independencia estadística. La situación planteada nos propone una población muy grande, de manera que si el Muestreo fuera con Reposición o Sin Reposición, con n (4) menor o igual al 5% del tamaño de la población (∞), las probabilidades permanecerán constantes a lo largo de todo el experimento. Es decir x ≈ B(n, P).

<u>Datos</u>:

(*Grupo* 0)	→	*P* (0)	= 0,46	
(*Grupo* A)	→	*P* (A)	= 0,40	
(*Grupo* B)	→	*P* (B)	= 0,11	
(*Grupo* AB)	→	*P* (AB)	= 003	$n = 4$

Si bien los datos no son dicotómicos, las preguntas en particular sí lo son:

a) En este caso, la población queda dicotomizada de la siguiente manera:

Grupo 0: x: éxito → $P = 0{,}46$
Otros Grupos: x': fracaso → $Q = 0{,}54$

<u>Entonces</u>:

$$\Pr(x = 2 \quad n = 4 \quad P = 0{,}46) = C_4^2 P^2 Q^{4-2} = \frac{4!}{2!\,2!}(0{,}46)^2(0{,}54)^2$$

$$= \frac{4 \times 3 \times 2 \times 1}{2 \times 1 \times 2 \times 1} \times 0{,}2116 \times 0{,}2916 = \underline{0{,}3702}$$

Nota aclaratoria:

$P = 0{,}46$ es la probabilidad de encontrar la categoría *Grupo* 0 en la población.

Pr = 0,3702 es la probabilidad de encontrar dos personas del *Grupo* 0 en una muestra de tamaño 4.

b) La población queda ahora dicotomizada, de la siguiente manera:

Grupo AB: x: éxito → $P = 0{,}03$
Otros Grupos: x': fracaso → $Q = 0{,}97$

<u>Entonces</u>:

$$\Pr(x = 0 \quad n = 4 \quad P = 0{,}03) = C_4^0 P^0 Q^4 = \frac{4!}{4!\,0!}(0{,}03)^0(0{,}97)^4 = 0{,}8853$$

c) x: Grupo A → $P = 0{,}40$

Entonces:

x_A	$P(x_A) = P(x = x_A \quad n = 4 \quad P = 0{,}40)$
0	0,1296
1	0,3456
2	0,3456
3	0,1536
4	0,0256
	Σ 1

d)

$x_A P(x_A)$	$x_A^2 P(x_A) = [x_A P(x_A)]x_A$
0	0
0,3456	0,3456
0,6912	1,3824
0,4608	1,3824
0,1024	0,4096
1,6	3,52

$$E(x_A) = \sum_{x_A=0}^{4} x_A P(x_A) = 1{,}6 \qquad \text{ó} \qquad E(x_A) = nP = 4 \times 0{,}40 = \underline{1{,}6}$$

$$V(x_A) = \sum_{x_A=0}^{4} x_A^2 P(x_A) - [E(x_A)]^2 = 3{,}52 - (1{,}6)^2 = 0{,}96 \qquad \text{ó}$$

$$V(x_A) = nPQ = 4 \times 0{,}40 \times 0{,}60 = \underline{0{,}96}$$

3) Se lanza diez veces un tetraedro equilibrado que tiene tres caras pintadas de azul y una de amarillo. Preguntas:

a- Cuál es la probabilidad que salgan siete caras pintadas de azul?

b- Cuál es la probabilidad que salgan más de seis caras pintadas de azul?

c- Cuál es la probabilidad que salgan cinco, seis, siete ú ocho caras pintadas de azul?

Datos:

$n = 10$

x: caras pintadas de azul $\qquad P = 3/4 = 0{,}75$

x': caras pintadas de amarillo $\qquad Q = 1/4 = 0{,}25$

a) $\Pr(x = 7 \quad n = 10 \quad P = 0{,}75) = \Pr(x = 3 \quad n = 10 \quad Q = 0{,}25) = 0{,}2503$


x $\qquad P = 0{,}75$


x' $\qquad P = 0{,}25$

$x' = n - x \quad \rightarrow \quad Q = 1 - P$

$3 = 10 - 7 \quad \rightarrow \quad 0{,}25 = 1 - 0{,}75$

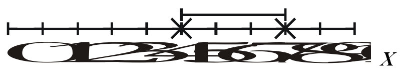
b) $\Pr(x > 6 \quad n = 10 \quad P = 0{,}75) = \Pr(x \geq 7 \quad n = 10 \quad P = 0{,}75) =$

$P = 0{,}75$

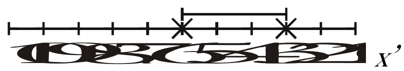
$1\text{-}P = 0{,}25$

$= \Pr(x' \leq 3 \quad n = 10 \quad 1 - P = 0{,}25) = 0{,}7759$

c) $\Pr(5 \leq x \leq 8 \quad n = 10 \quad P = 0{,}75) =$



$P = 0{,}75$



$P = 0{,}25$

$= \Pr(2 \leq x \leq 5 \quad n = 10 \quad 1 - P = 0{,}25) =$

$= \Pr(x' \leq 5 \quad n = 10 \quad 1 - P = 0{,}25) - \Pr(x' \leq 1 \quad n = 10 \quad 1 - P = 0{,}25) =$

$= 0{,}9803 - 0{,}2440 = 0{,}7363$

1.3. <u>MODELO HIPERGEOMÉTRICO:</u> Número De Éxitos En n Pruebas

Características

La Distribución Binomial se basa en el supuesto de que la población es finita ó infinita y de que la muestra aleatoria se toma con reemplazo, de manera que las observaciones posibles sean independientes entre sí. La probabilidad permanecerá así invariable para toda la observación sucesiva. Esto ocurre también cuando la población es infinita y el muestreo es sin reposición, siempre que $n \leq 5$ % de N.

Cuando la población es finita, y la muestra aleatoria se toma sin reemplazo, la probabilidad cambiará para cada nueva observación, con lo cual habrá dependencia estadística.

En tales circunstancias se tendrá una distribución de probabilidad que se llama *Distribución Hipergeométrica.*

Función de probabilidad: Función de Cuantía

En la Distribución Binomial, hemos visto que la probabilidad es constante, pues se repone el elemento extraído en cada muestra. Supongamos ahora, que a partir de una población madre finita, extraemos una muestra de tamaño n sin reponer el elemento extraído en cada prueba.

Ello equivale, indudablemente, a extraer de la población considerada los n elementos que conforman la muestra en forma simultánea. Se trata de determinar una función que nos permita calcular la probabilidad de que en dicha muestra obtengamos un numero x de éxitos, o sea x elementos con la propiedad o característica que hemos designado como éxito.

Por ejemplo, si tenemos un lote de 500 facturas y el 10% tienen errores de cálculo, al tomar una muestra al azar, pero simultáneamente, de 20 facturas, ¿cuál es la probabilidad dé encontrar *5* facturas con errores de cálculo?

Generalizando el problema, supongamos tener una población finita, de N elementos dentro de los cuales sabemos que X de ellos tienen una cierta propiedad o característica (A), y los restantes $N - X$, no la tienen ($\overline{A}$).

A	$\overline{A}$	N
X	$N - X$	

Extraemos una muestra de tamaño n y se pregunta: ¿cuál es la probabilidad de encontrar en dicha muestra x elementos que pertenezcan a la clase A?

La probabilidad buscada se simboliza por: P($x = x$; N; X; n)

Para encontrar la probabilidad buscada nos basaremos en la definición clásica de probabilidad:

$$\frac{\text{Casos favorables}}{\text{Casos posibles}}$$

Casos favorables: se trata de un evento compuesto por varios eventos mutuamente excluyentes y su número está dado por el producto:

$$C_X^x \; C_{N-X}^{n-x}$$

El primer factor C_X^x son las distintas formas que pueden presentarse x elementos con la propiedad A (éxitos), dentro del total que tienen dicha propiedad, o sea los X; pero siendo la muestra de tamaño n, deberán presentarse junto con los x éxitos, $n - x$ fracasos, o sea $n - x$ elementos de la clase $\overline{A}$, cuyo total es $N - X$, y dicho número está dado por el segundo factor

$$C_{N-X}^{n-x}$$

Finalmente, como cada grupo de los x éxitos puede combinarse con cada uno de los grupos de los n–x fracasos, el total de casos está dado por el producto de ambos.

Casos posibles: el total de muestras de tamaño n que podemos extraer sin reposición de una población N, es C_N^n

Por lo tanto, la función de probabilidad buscada es:

$$P(X = x;\ N;\ X;\ n) = \frac{C_X^x C_{N-X}^{n-x}}{C_N^n}$$

Y es una función de probabilidad o cuantía, que depende de tres parámetros:

N: Población dicotomizada

n: Tamaño de la muestra

X: Número o parte de la población que tiene determinada propiedad o característica.

Por otro lado, *x* es la variable real que indicará los valores esperados de la variable aleatoria.

En el ejemplo dado:

N = 500, de los cua1es X =10% = 50, luego: *N* – X = 450

n = 20, de los cuales *x* = 5, luego *n* – *x* = 15

Entonces:

$$P(X = 5, N = 500, X = 50, n = 20) = \frac{C_{50}^5\ C_{450}^{15}}{C_{500}^{20}}$$

Función de Distribución o Función de Acumulación

$$P(X \leq x;\ N;\ X;\ n) = \sum \frac{C_X^x C_{N-X}^{n-x}}{C_N^n}$$

Parámetros

Esperanza

Para derivar la esperanza de una variable con Distribución Hipergeométrica, observemos que la variable *x*, Hipergeométrica, también puede ser considerada como una suma, de igual modo que en el caso Binomial, de *n* variables y_i (Bipuntuales), excepto que en el Modelo Hipergeométrico y_1, y_2,..., y_n son dependientes.

Pero como la propiedad aditiva de la esperanza no requiere independencia de las variables (sean o no independientes), es cierto aún que $E(x) = E(y_1) + \cdots + E(y_n)$, donde cada $E(y_1)$ es la probabilidad de *y* en la i-ésima prueba (*X*/ *N*), si no se sabe lo que ha sucedido en pruebas anteriores.

Así, la esperanza de la variable Hipergeométrica *x*, coincide con la de la variable Binomial.

$$E(x) = nP \quad , \text{ donde } P = \frac{X}{N}$$

Entonces $$E(x) = n\ \frac{X}{N}$$

Varianza

La varianza no es aditiva para variables dependientes. Pero puede demostrarse que en el caso Hipergeométrico la varianza tiene la siguiente expresión:

$$V(x) = \left(\frac{N-n}{N-1}\right) nPQ \qquad \text{Donde} \quad P = \frac{X}{N}$$

$$\text{y} \qquad Q = 1 - P = 1 - \frac{X}{N}$$

Entonces

$$V(x) = \frac{N-n}{N-1} n \frac{X}{N}\left(1 - \frac{X}{N}\right) = \frac{N-n}{N-1} n \frac{X}{N}\left(\frac{N-X}{N}\right)$$

Vemos que también el modelo binomial es apropiado y $V(x) = nPQ = n\left(\frac{X}{N}\right)\left(1 - \frac{X}{N}\right)$. Cuando se supone muestreo sin reposición, como el Modelo Hipergeométrico, se introduce un factor de corrección de poblaciones finitas $\frac{N-n}{N-1}$, que cuando $N \to \infty$ tiende a 1.

O sea que en el muestreo sin reemplazo, cuando la población es suficientemente grande, la varianza tiende a ser igual que la varianza de la variable Binomial y ello permite trabajar como si la muestra fuera con reposición y aplicar directamente la Distribución Binomial.

Entonces, cuando el tamaño de la muestra es grande comparado con el tamaño de la población y el muestreo es sin reposición, partiendo de una población finita, el Modelo Hipergeométrico es de aplicación.

Desviación Estándar

$$\sigma_X = \sqrt{\frac{N-n}{N-1} n \frac{X}{N}\left(\frac{N-X}{N}\right)}$$

Tablas Usuales

Existe una tabla general de probabilidades: *Tabla V*, que contiene tanto la *Función de Probabilidad* ($p(x)$, según la tabla) como la *Función de Acumulación* ($P(r)$, según la tabla).

Al usarla, debe observarse una propiedad interesante de la Distribución Hipergeométrica: cuando X y n se intercambian, no hay efectos en las probabilidades.

<u>Ejemplo</u>:

Si se extraen juntas al azar 4 fichas de una urna que contiene 6 fichas negras y 4 blancas, ¿Cuál es la probabilidad de que 2 ó menos fichas negras sean extraídas?

Datos:

$N = 10$
$X = 6$
$n = 4$

Entonces:

$$\Pr(x \leq 2,\ N = 10,\ n = 4,\ X = 6)$$

Si queremos utilizar la tabla para la resolución de dicha probabilidad, veremos que el valor $k = 6$, no existe, por lo tanto deberemos resolver de la siguiente manera:

$$\Pr(x \leq 2,\ N = 10,\ n = 4,\ X = 6) = \sum_{x=0}^{2} \frac{C_6^x C_4^{4-x}}{C_{10}^4} = \frac{C_6^0 C_4^4}{C_{10}^4} + \frac{C_6^1 C_4^3}{C_{10}^4} + \frac{C_6^2 C_4^2}{C_{10}^4} = \underline{0{,}547619}$$

Ahora bien, si aplicamos la propiedad, e intercambiamos los valores de X y n, obtenemos la probabilidad directamente de la tabla:

$$\Pr(x \leq 2,\ N = 10,\ n = 6,\ X = 4) = \underline{0{,}547619}$$

Como se observa, se ha llegado a idénticos resultados.

Ahora, si se extraen 5 fichas sin reposición, de la misma urna, ¿cuál es la probabilidad de que: 1) Sean todas negras, 2) Cuatro o más sean negras?

1) $\Pr(x = 5,\ N = 10,\ n = 5,\ X = 6) = \Pr(x = 5,\ N = 10,\ n = 6,\ X = 5) = \underline{0{,}023810}$

2) $\Pr(x \geq 4,\ N = 10,\ n = 5,\ X = 6) = \Pr(x \geq 4,\ N = 10,\ n = 6,\ X = 5) =$
$= 1 - \Pr(x \leq 3,\ N = 10,\ n = 6,\ X = 5) = = 1 - 0{,}738095 = \underline{0{,}261905}$

Nota: Obsérvese que:

$$x = \left| \begin{array}{ll} 0,1,2,\ldots,n & \text{si } n \leq X \\ 0,1,2,\ldots,X & \text{si } n > X \end{array} \right.$$

1.4. Modelo Binomial y Modelo Hipergeométrico: Proporción De Éxitos En n Pruebas

Hasta el momento hemos tratado el número de éxitos en n pruebas. Pero es frecuente, que lo que interesa no es el número de éxitos, sino la proporción de los mismos, como por ejemplo estudiar la proporción de defectuosos en un lote en vez del número de defectuosos.

Así, tendremos:

a) $P = \frac{X}{N}$, que simboliza y define a la *Proporción Poblacional*, donde *X* es el número de éxitos en la población y *N* el tamaño poblacional.

b) $p = \hat{P} = \frac{x}{n} = \frac{\sum_{i=1}^{n} y_i}{n}$, que simboliza y define a la *Proporción Muestral,* donde *x* es el número de éxitos en pruebas, *n* el tamaño muestral y "*y*" la variable aleatoria Bernoulli.

Como se trata de una proporción, los valores de $\hat{P}$ son por lo general valores fraccionarios o decimales como: $0, \frac{1}{n}, \frac{2}{n}, \ldots, 1$.

El valor de $\hat{P}$ depende de los valores de *x* y *n*. Dado el valor de *n*, entonces $\hat{P}$ está completamente determinado si se conoce *x*. Así, la probabilidad de que ocurra una proporción dada de éxitos es idéntica a la probabilidad de que se de este número de éxitos.

Como $\hat{P}$ es un estadígrafo muestral (medida calculada en base a datos de muestra), su distribución es una distribución muestral.

La distribución por muestreo de $\hat{P}$ toma dos formas distintas, dependiendo de si la población muestreada es finita o infinita.

Recordemos que una muestra aleatoria se puede tomar con o sin reemplazo:

1- Si la muestra se toma con reemplazo de una población finita o infinita, la unidad tomada se vuelve a dejar en la población y el número de unidades disponibles para seguir la operación no se afecta.

Esto también es cierto cuando la muestra se toma de una población infinita sin reemplazo, es decir cuando la unidad escogida no se vuelve a la población, siempre que n< 5% s/N.

2- Cuando se toma un elemento sin reemplazar de una población finita, el número de unidades que queda tras cada unidad que se saca se reduce en una unidad, y en consecuencia la probabilidad de sacar cualquier unidad restante en operaciones sucesivas se aumenta.

Esto también es cierto cuando la muestra se toma de una población infinita sin reemplazo, es decir cuando la unidad escogida no se vuelve a la población, siempre que n≥5% s/N.

Veamos entonces los dos casos:

1- Si se obtiene una muestra al azar de *n* elementos, *con reposición*, donde *x* es binomial, la distribución por muestreo de *p* obedece a la ley de probabilidad binomial, y podemos escribir:

Función de Probabilidad: Función de cuantía

$$\Pr\left(\hat{P} \leq \hat{P}_i\right) = \Pr\left(\hat{P} = \frac{x}{n}\right) = \Pr\left(x = x_i\right) = \begin{vmatrix} C_n^x P^x (1-P)^{n-x} & x = 0,1,2,\ldots,n \\ 0 & \forall \text{ otro } x \end{vmatrix}$$

Entonces la probabilidad de que la proporción muestral, es decir, la probabilidad de que se presente una proporción de éxitos dados en la muestra, es igual a la probabilidad de que se presente ese número de éxitos en la muestra.

Función de Distribución o Función de Acumulación

$$\Pr\left(\hat{P} \leq \hat{P}_i\right) = \Pr\left(x \leq x_i\right) = \sum_{x=0}^{x_i} C_n^x P^x (1-P)^{n-x}$$

Parámetros

Esperanza

$$E\left(\hat{P}\right) = E\left(\frac{x}{n}\right) \qquad \text{pues } \hat{P} = \frac{x}{n}$$

$$= \frac{1}{n} E(x) =$$

la esperanza del producto de una constante y variable, es igual al producto de la constante por la esperanza de la variable.

$$= \frac{1}{n} nP =$$

pues $E(x) = nP$, ya que x es una variable binomial.

Entonces: $\boxed{E\left(\hat{P}\right) = P}$

Varianza

$$V\left(\hat{P}\right) = V\left(\frac{x}{n}\right) = \frac{1}{n^2} V(x)$$

$$= \frac{1}{n^2} nP(1-P) =$$

$$= \frac{P(1-P)}{n}$$

La varianza del producto de una constante y una variable es igual al producto del cuadrado de la constante por la varianza de la variable.
Pues V(x) = n P(1–P), ya que x es una variable Binomial

Entonces: $\boxed{V\left(\hat{P}\right) = \frac{P\,(1-P)}{n}}$

Desviación Estándar

$$\sigma_{\hat{P}} = \sqrt{\frac{P(1-P)}{n}}$$

Si se obtiene una muestra al azar, de n elementos, *Sin Reposición*, donde x es Hipergeométrica, la distribución de p obedece a la ley de probabilidad Hipergeométrica, y podemos escribir:

Función de Probabilidad: Función de cuantía

$$P\left(\hat{P} = \hat{P}_i\right) = P\left(\hat{P} = \frac{x}{n}\right) = \Pr\left(x = x_i\right) = \frac{C_X^x \; C_{N-X}^{n-x}}{C_N^n}$$

Función de Distribución o de Acumulación

$$P\left(\hat{P} \leq \hat{P}_i\right) = P\left(\hat{P} \leq \frac{x}{n}\right) = \Pr\left(x \leq x_i\right) = \sum_{x=0}^{x_i} \frac{C_X^x \; C_{N-X}^{n-x}}{C_N^n}$$

Parámetros

En este caso, $E\left(\hat{P}\right)$ es todavía idéntica a *P*, pero la varianza debe ser ajustada por un factor de corrección de población finita (FCPF) igual a: $\frac{N-n}{N-1}$

Entonces:

$$E\left(\hat{P}\right) = P$$

$$V\left(\hat{P}\right) = \frac{P(1-P)}{n} \frac{N-n}{N-1}$$

$$\sigma_{\hat{P}} = \sqrt{\frac{P(1-P)}{n} \frac{N-n}{N-1}}$$

Si la muestra es pequeña con relación a la población, por ejemplo, la primera es menor del 5% de la última, el FCPF se aproximará a la unidad. En consecuencia la introducción del FCPF producirá una diferencia prácticamente despreciable. En otras palabras, si se toma una muestra pequeña de una población grande, podemos tratar las observaciones de la muestra como si fueran variables independientes y puede calcularse sin el FCPF.

<u>Ejemplo</u>:

Ahora veremos un ejemplo sencillo para demostrar cómo se deriva una Distribución de una Proporción Muestral, y comprobar que la esperanza de $\hat{P}$ es idéntica a P y que la varianza es definida por:

$$\sigma_{\hat{P}}^2 = \frac{P(1-P)}{n} \qquad \text{En M.C.R.}$$

$$\sigma_{\hat{P}}^2 = \frac{P(1-P)}{n} \frac{N-n}{N-1} \qquad \text{En M.S.R.}$$

Supóngase que el gerente de una compañía se interesa en conocer la proporción de empleados menores de 35 años y se propone tomar una muestra de los archivos de la compañía para obtener información a este respecto. Para simplificar el problema supongamos que hay tan sólo cinco empleados en la compañía, en otras palabras que el tamaño de la población es N=5 y se ha de elegir una muestra de 3 empleados.

Las edades de los 5 empleados están dadas en la siguiente tabla:

Empleados	Edad en años (x)
A	27
B	39
C	30
D	36
E	42

NOTA: *Evidentemente, en la práctica no tomaríamos una muestra de una población de 5 individuos, sino que realizaríamos un censo.*

Resolución:

Datos: N=5 n = 3 X= 2 (cantidad de individuos que responden a la característica analizada: empleados menores de 35 años).

P = 2/5 = 0,40

1- Con reposición

x_i	$\hat{P}_i$	$P(\hat{P}_i) = \Pr(x = x_i \quad n = 3 \quad P = 0{,}40)$	$\hat{P}_i P(\hat{P}_i)$	$\hat{P}_i^2 P(\hat{P}_i) = \left[\hat{P}_i P(\hat{P}_i)\right]\hat{P}_i$
0	0	0,2160	0	0
1	1/3	0,4320	0,144	0,048
2	2/3	0,2880	0,192	0,1280
3	3/3	0,0640	0,064	0,064
		$\sum = 1$	$\sum = 0{,}40$	0,24

En una muestra de tamaño 3, pueden presentarse cuatro valores posibles para la variable: 0 (no aparezca ningún empleado menor de 35 años), 1 (aparezca un empleado menor de 35 años) y así sucesivamente. Las probabilidades fueron encontradas en la tabla I de la Distribución Binomial, pues trabajamos con reposición.

Podríamos calcular cualquier probabilidad puntual o acumulada, sin necesidad de construir la tabla, por ejemplo si se nos pregunta cuál es la probabilidad de que p =1/3?, entonces:

Pr ($\hat{P}_i$ = 1/3) = Pr(x = 1 n = 3 P = 0,40)

Ahora calculamos:

$$E(\hat{P}) = \sum_{i=1}^{4} \hat{P}_i P(\hat{P}_i) = 0{,}40 \Rightarrow E(\hat{P}) = P$$

$$V(\hat{P}) = \sum_{i=1}^{4} \hat{P}_i^2 P(\hat{P}_i) - \left[E(\hat{P}_i)\right]^2 = 0{,}24 - (0{,}40)^2 = 0{,}08$$

$$V(\hat{P}) = \frac{P(1-P)}{n} = \frac{(0{,}40)(0{,}60)}{3} = 0{,}08$$

2- *Sin Reposición*

Siendo $n = 3$ y $k = 2$, x asumirá los valores 0,1 y 2.
Luego:

x_i	$\hat{P}_i$	$P(\hat{P}_i) = \Pr(x = x_i \quad N = 5 \quad X = 2 \quad n = 3) = C_2^x C_3^{3-x} / C_5^3$	$\hat{P}_i P(\hat{P}_i)$	$\hat{P}_i^2 P(\hat{P}_i)$
0	0	0,10	0	0
1	1/3	0,60	0,20	0,0667
2	2/3	0,30	0,20	0,1333
		$\sum \;= 1$	0,40	0,20

Las probabilidades fueron encontradas en la tabla de la Distribución Hipergeométrica, pues trabajamos sin reposición.

Calculamos ahora:

$$E(\hat{P}) = \sum_{i=1}^{4} \hat{P}_i P(\hat{P}_i) = 0{,}40 \;\Rightarrow\; E(\hat{P}) = P$$

$$V(\hat{P}) = \sum_{i=1}^{4} \hat{P}_i^2 P(\hat{P}_i) - \left[E(\hat{P}_i)\right]^2 = 0{,}20 - (0{,}40)^2 = 0{,}04$$

$$V(\hat{P}) = \frac{P(1-P)}{n}\frac{N-n}{N-1} = \frac{(0{,}4)(0{,}6)}{3}\frac{5-3}{5-1} = 0{,}08 \times 0{,}5 = 0{,}04$$

Configuración

Para cualquier n:	Si	$\hat{P} = 0{,}5$	$\Rightarrow$ Simetría
Para un n dado:	Si	$\hat{P} > 0{,}5$	$\Rightarrow$ Asimetría Izquierda
	Si	$\hat{P} < 0{,}5$	$\Rightarrow$ Asimetría Derecha.

Sin embargo, conforme el tamaño de la muestra se hace mayor, la distribución de $\hat{P}$ se hace más y más simétrica para cualquier valor de la proporción de población P. En efecto, para tamaños de muestra razonablemente grandes, la distribución de muestreo de $\hat{P}$ es más o menos simétrica.

1.5. ***MODELO DE POISSON***

Características

Supóngase que se desea encontrar la distribución de probabilidad del número de accidentes automovilísticos en un cruce en particular durante un período de una semana.

A primera vista esta variable aleatoria, número de accidentes, puede no parecer relacionada, ni remotamente, a una variable aleatoria binomial, pero veremos que existe un vínculo importante.

Considere que el período de una semana del ejemplo anterior se divide en n subintervalos, cada uno de los cuales es tan pequeño que podría ocurrir en él a lo más un accidente, con una probabilidad diferente de 0. Si se denota la probabilidad de un accidente en cualquier intervalo como P, se tiene entonces, para fines prácticos:

Pr (ningún accidente en un subintervalo) = 1-P

Pr (un accidente en un subintervalo) = P

y

Pr (más de un accidente en un subintervalo) = 0

De este modo el número total de accidentes en una semana es exactamente el número total de subintervalos que contienen un accidente. Si se puede considerar la ocurrencia de accidentes como independiente de un intervalo a otro, el número total de accidentes tiene una distribución binomial.

Aunque no hay una manera única de elegir los subintervalos, y por eso no conocemos n ni P, parece razonable que la probabilidad p de un accidente en uno de los subintervalos decrecerá al dividir la semana en un número n cada vez mayor de subintervalos.

Haciendo λ = nP y tomando límite de la probabilidad binomial cuando n → ∞

$$P(x = x_i,\ \mu = \lambda) = \left| \begin{array}{ll} \dfrac{e^{-\lambda}\lambda^x}{x!} & ; \quad x = 0,1,2,\ldots,\infty \\ 0 & \quad \forall \text{ otro } x \end{array} \right.$$

Las variables aleatorias que tienen esta distribución se denominan variables aleatorias de Poisson.

Muchos hechos no ocurren como resultado de un número definido de pruebas de un experimento, sino para el número X de eventos raros que ocurren infrecuentemente en puntos de tiempo, espacio o volumen o cualquier otra dimensión al azar, tales como:

- La demanda de servicio por unidad de tiempo a una cajera o a una vendedora de una tienda, a un empleado de una fábrica, a una garita de peaje, a una instalación de manejo de carga de un puerto, a las líneas troncales de un intercambio telefónico, a un técnico de servicio de un taller de maquinaria.

- El número de defectos en la superficie de una mesa, en una unidad de cristalería o en una pieza de tejido.
- El número de blancos de bombardeo por milla cuadrada.
- El número de glóbulos rojos en una muestra de sangre.

En estos casos, la distribución de Poisson proporciona un buen modelo para la distribución de probabilidad de estos hechos.

Para cada una de estas variables aleatorias mencionadas antes los siguientes postulados son apropiados a veces:

1- El número de ocurrencias del hecho es independiente de una unidad (intervalo de tiempo, espacio o volumen) especificada a otra.
2- El valor esperado de la variable es proporcional al tamaño de la unidad especificada.
3- La probabilidad de más de una ocurrencia del hecho en una unidad especificada muy pequeña es despreciable en comparación con la probabilidad de una sola ocurrencia; por lo tanto puede despreciarse.

En estas condiciones, se define el modelo Poisson cuya distribución de probabilidades da la probabilidad del número de ocurrencias por unidad especificada, y es definida completamente por su promedio de ocurrencias por unidad especificada, λ; su único parámetro.

Es decir, surgen de procesos en los que hay una observación por unidad de tiempo, espacio o volumen, que se caracterizan por el número de éxitos esperados en la unidad específica, tal como la Distribución Binomial se caracteriza por el número de éxitos en n pruebas.

Lo antedicho de "por unidad de tiempo, espacio o volumen", equivale al tamaño n de la muestra en una Distribución Binomial.

Ya hemos visto que la Distribución Binomial se basa en el proceso de Bernoulli, caracterizado por su parte, por la probabilidad de un éxito en una prueba. ¿Qué es entonces lo que constituye una prueba en un proceso de Poisson? En el caso de un lote de partes, la observación de una sola parte para determinar si es o no defectuosa es, naturalmente, una prueba.

Pero, en cuanto a las llamadas telefónicas recibidas por minuto?. Si se divide el minuto en 60 segundos, hay todavía la posibilidad de recibir más de una llamada en un segundo dado. Para que esto no ocurra, hay que dividir el minuto en subdivisiones mucho menores, tales como 1/100 de segundo. Se podrá confiar entonces en que la probabilidad de dos o más llamadas en un intervalo de tiempo tan insignificante es tan pequeña que se la puede dejar de lado para fines prácticos. Cuando se divide una unidad de tiempo o espacio en muchas subdivisiones (n grande), y la probabilidad de éxito en una sola prueba es pequeña (probabilidad pequeña), se estará en una situación cercana a la de una Distribución de Poisson, considerándola el límite de la Distribución Binomial.

Sea $\lambda = \mu = nP$, la media del número esperado de éxitos en un tiempo o espacio dados, y designe con la variable x, al número de éxitos en una muestra de tamaño n con los posibles valores (0, 1, 2, 3,...), entonces:

***Función de Probabilidad*: Función de Cuantía**

La función de probabilidad o función de cuantía, estará dada por la fórmula:

$$P(x = x_i,\, \mu = \lambda) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!} & ; \quad x = 0,1,2,\ldots,\infty \\ 0 & \forall \text{ otro } x \end{cases}$$

Donde e es una constante igual a 2,71828 y x es cualquier valor posible que puede tomar X.

Función de Distribución o Función de Acumulación

Su función de Distribución o de acumulación es:

$$F(x) = P(x \le x_i;\, \mu = \lambda) = \sum_{x=0}^{x_i} \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 1,2,3,\ldots,\infty$$

Parámetros

Esperanza $E(x) = nP = \lambda = \mu$

Varianza $V(x) = nP = \lambda = \mu$

Desviación Estándar $DS(x) = \sqrt{nP} = \sqrt{\lambda} = \sqrt{\mu}$

La varianza de la distribución Binomial es nP (1 – P). Al tender P hacia 0, el límite del último factor 1 – P es 1, y la varianza tiende hacia nP o μ.

La Distribución de Poisson tiene solo un parámetro que es $\lambda = \mu$, que depende del tamaño de la unidad especificada, y que λ cambia proporcionalmente al cambiar la unidad especificada, de manera que duplicando la unidad especificada también se duplica el número esperado de ocurrencias por nueva unidad especificada.

En un ejemplo de llamada telefónica en que λ = 4 llamadas por minuto y nos preguntamos ¿Cuál es la probabilidad de que no ocurran más de dos llamadas durante los dos minutos siguientes?, ahora λ=4(2) =8 llamadas en dos minutos y Pr (x=2 λ=8).

Del mismo modo si nos preguntáramos, ¿Cuál es la probabilidad de que no se produzca ninguna llamada en cualquier intervalo de 30 segundos?, λ=4/2 llamadas en 30 segundos y Pr (x=0, λ=2)

Configuración

Es de asimetría positiva, sin embargo a medida que el tamaño de la muestra aumenta tiende a la simetría, o bien el valor de λ, su expectativa.

Tablas Usuales

Al igual que para la Distribución Binomial contamos con dos tablas:

a. *Tabla III*, referida a la Función de Probabilidad, o sea *Pr(x = x_i;λ).*

Las entradas de la tabla son valores de $\frac{e^{-\lambda}\lambda^{x}}{x!}$ para los valores de *x* y *n* dados.

<u>Ejemplo</u>:
Pr (x = 3; λ = 5) = 0,1404

Para responder al planteo utilizamos la Tabla III, donde ubicamos previamente el valor de *λ ó m* que en nuestro caso es 5. Debajo de este valor aparecen un grupo de probabilidades para cada valor de *x;* el resultado es entonces, el valor de probabilidad que corresponde a x = 3 (0,1404).

b. *Tabla IV*, referida a la *Función de Acumulación*, o sea *Pr(x ≤ x_i;λ).*

Las entradas en la tabla son valores de $\frac{\sum_{x=0}^{x_i} e^{-\lambda}\lambda^{x}}{x!}$ para los valores indicados de x_i y λ, es

decir acumula para valores ≤ x_i (desde 0 hasta x_i dado).

<u>Ejemplos</u>:

a) Pr (x ≥ 2; λ = 4)

A esta probabilidad no la encontraremos directamente en la tabla IV, entonces hacemos:

$$\Pr(x \geq 2; \lambda = 4) = 1 - \Pr(x \leq 1; \lambda = 4) = 1 - 0{,}0916 = 0{,}9084$$

b) Pr (x ≤ 1; λ = 4) = 0,0916

La encontramos directamente en la tabla para λ = 4 buscamos x =1.

<u>Ejercicios</u>

1- Si el 3‰ de las facturas confeccionadas por un empleado son defectuosas hallar la probabilidad de que en una muestra de 1000 facturas

(a) ninguna
(b) 2
(c) 5
(d) más de 5,
(e) entre 3 y 8, ambos inclusive
(f) 2 o menos, sean defectuosas (con algún error).

De los datos surge que: $P = 0{,}003$ y $n = 1000$, luego $nP = 3$

a) Pr (x = 0 λ = 3) = <u>0,0498</u>

b) Pr (x = 2 λ = 3) = <u>0,2240</u>

c) Pr (x = 5 λ = 3) = <u>0,1008</u>

d) Pr (x > 5 λ = 3) = Pr ($X \geq 6$ λ = 3) = 1 – Pr ($x \leq 5$ λ = 3)== 1– 0,9161 = <u>0,0839</u>

e) Pr ($3 \leq x \leq 8$, λ = 3) = Pr ($x \leq 8$, λ=3) – Pr ($x \leq 2$, λ =3)= =0,9962 – 0,4232 = <u>0,5730</u>

f) Pr ($x \leq 2$, λ = 3) = <u>0,4232</u>

2- Cierta ciudad tiene, en promedio, 12 muertes de tráfico cada tres meses ¿Cuál es la probabilidad de que en cualquier mes haya:
a) Más de cuatro muertes de tráfico?
b) Exactamente cuatro muertes?

De los datos surge que λ = 4 por mes,

a) Pr (x > 4 ; λ = 4) = 0,3712

b) Pr (x = 4; λ = 4) = 0,1954

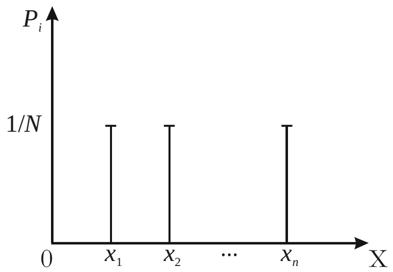
1.6. *MODELO UNIFORME DISCRETO*

Una distribución de probabilidades es uniforme o rectangular, cuando la probabilidad asociada con todos y cada uno de los resultados es una constante.

La función de probabilidad de este modelo puede escribirse:

$$P(x = x_i) = \left| \begin{array}{l} P_i = 1/N \\ 0 \quad \forall \text{ otro } i \end{array} \right.$$

Gráficamente:

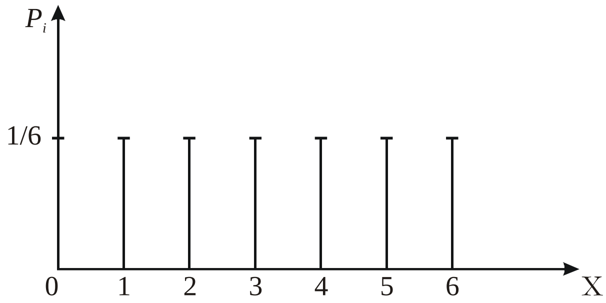


En general, se aplica a un experimento con N resultados mutuamente exclusivos e igualmente probables, para lo cual debe utilizarse una tabla de dígitos al azar que trataremos más adelante.

Hemos encontrado antes el modelo de probabilidad uniforme discreto con la distribución de probabilidades del número de puntos en un dado perfecto pues la probabilidad asociada con todos y cada uno de los resultados era una constante igual a 1/6 y la función de probabilidad:

$$P(x = x_i) = \begin{vmatrix} 1/6 & \text{para } x = 1,2,3,4,5,6 \\ 0 & \forall \text{ otro } i \end{vmatrix}$$

Gráficamente:



Hasta aquí hemos tratado las distribuciones discretas de probabilidad, es decir distribuciones para variables discretas, o sea aquéllas que asumen valores enteros o fraccionarios especificados.

En estos casos, la probabilidad en un punto tiene valor, por lo cual son sumamente importantes las expresiones ≤ o ≥ que se diferencian de < o >, ya que estas últimas no incluyen los puntos mencionados.

<u>Ejemplo</u>:

Si planteamos

Pr (x <3), esta expresión no incluye al 3, es decir es equivalente a

Pr (x ≤ 2).

Variables Aleatorias Discretas

Base de los Modelos: Distribución Bipuntual (n = 1)

Resultados	y	P(y)
Fallo	0	$Q = 1-P$
Éxito	1	P
		$1=P+Q$

$E(y)=P$
$V(y)=PQ$

Variables (1)	***Modelo (2)***	***Expresión Simbólica de (2)***	***Valores Posibles***	***Características***
x *Número de Éxitos en n Pruebas*	***Binomial***	$x \sim \beta(n,P)$	$x = 0,1,2,...,n$	*Independencia Estadística: Las Probabilidades permanecen Constantes a lo largo De todo el Experimento* ***Población Infinita*** MCR MSR: n ≤ 5% s/N ***Población Finita*** MCR
		$\hat{P} \sim \beta(n,P)$	$\hat{P} = 0, \frac{1}{n}, \frac{2}{n}, ..., \frac{n}{n}$	
p = $\hat{P} = \frac{x}{n}$ *Proporción de Éxitos en n Pruebas*	***Hipergeométrico***	$x \sim H(N,X,n)$	$x = \begin{cases} 0,1,2,...,X & n > X \\ 0,1,2,...,n & n \leq X \end{cases}$	*Dependencia Estadística: Las Probabilidades NO permanecen Constantes a lo largo De todo el Experimento* ***Población Infinita*** MSR MSR: n >5% s/N ***Población Finita*** MSR
		$\hat{P} \sim H(N,X,n)$	$\hat{P} = \begin{cases} 0, \frac{1}{n}, \frac{2}{n}, ..., \frac{X}{n} & n > X \\ 0, \frac{1}{n}, \frac{2}{n}, ..., \frac{n}{n} & n \leq X \end{cases}$	
$x \sim \beta(n,P)$	***Poisson***	$x \sim P(\lambda = nP)$	$x = 0,1,2,...,\infty$	*N Grande* *P Pequeño*

Modelos Especiales de Probabilidad para Variables Aleatorias Discretas

Variable	*Distribución*	*Función De Probabilidad*	*Función De Acumulación*	E( )	V( )	DS( )
(γ)	Bipuntual	$P(\gamma=\gamma)=\left\lvert\begin{array}{l}P^{\gamma}(1-P)^{1-\gamma}\\0\end{array}\right.$		P	PQ	$\sqrt{PQ}$
(x)	Binominal	$P(X=x)=\left\lvert\begin{array}{l}C_n^x P^x(1-P)^{n-x}\\0\end{array}\right.$	$P(X\le x_i,n,P)=\sum_{x=0}^{x_i}C_n^x P^x(1-P)^{n-x}$	$n\,P$	$n\,P\,Q$	$\sqrt{nPQ}$
(x)	Hipergeométrica	$P(X=x,N,n,k)=\frac{C_k^x C_{N-k}^{n-x}}{C_N^n}$	$P(X\le x_i,N,n,k)=\sum_{x=0}^{x_i}\frac{C_k^x C_{N-k}^{n-x}}{C_N^n}$	$n\frac{k}{N}$	$\frac{N-n}{N-1}n\frac{k}{N}\left(1-\frac{k}{N}\right)$	$\sqrt{\frac{N-n}{N-1}n\frac{k}{N}\left(1-\frac{k}{N}\right)}$
($\hat{P}$)	Binomial	$P(\hat{P}=\hat{P}_i)=P(X=x,n,P)$	$P(\hat{P}\le\hat{P}_i)=P(X\le x_i,n,P)$	P	$\frac{P(1-P)}{n}$	$\sqrt{\frac{P(1-P)}{n}}$
($\hat{P}$)	Hipergeométrica	$P(\hat{P}=\hat{P}_i)=P(X=x,N,n,k)$	$P(\hat{P}\le\hat{P}_i)=P(X\le x_i,N,n,k)$	P	$\frac{P(1-P)}{n}\frac{N-n}{N-1}$	$\sqrt{\frac{P(1-P)}{n}\frac{N-n}{N-1}}$
(x)	Poisson	$P(X=x,\lambda)=\frac{e^{-\lambda}\lambda^x}{x!}$	$P(X\le x_i,\lambda)=\sum_{x=0}^{x_i}\frac{e^{-\lambda}\lambda^x}{x!}$	$n\,P$	$n\,P$	$\sqrt{nP}$

CAPITULO Nº 7: Modelos Especiales de Probabilidad
Variables Aleatorias Continuas

Objetivos Específicos

Que el estudiante:

Conozca *alguno de los diferentes modelos de probabilidad de variables aleatorias continuas.*

Distinga *las características de cada modelo, a lo fines de aplicarlos a situaciones reales.*

Reconozca *las funciones de densidad y de acumulación para el cálculo de probabilidades.*

Conozca *el manejo de las tablas usuales y como las mismas fueron calculadas a través de sus respectivas funciones de probabilidad y de acumulación.*

Conozca y comprenda *la importancia de la Ley de los Grandes Números y el Teorema Central del Límite*

Conozca y maneje tablas *de la Distribución Chi cuadrado, útil para la estimación y dócima de la varianza.*

Conozca y maneje *tablas de la Distribución t de Student, útil para la estimación y dócima de la media poblacional cuando se desconoce la varianza poblacional.*

Contenidos

MODELOS ESPECIALES DE PROBABILIDAD
VARIABLES ALEATORIAS CONTINUAS

También en esta Unidad es aplicable lo mencionado en Modelos Especiales de Probabilidad para Variables Discretas, en relación a que son modelos matemáticos apropiados para situaciones del mundo real en condiciones específicas. Son importantes porque ayudan a predecir la conducta de futuras repeticiones de un experimento, en otras palabras, un modelo probabilístico es una expresión matemática, resultante de un cúmulo de supuestos que tiene como propósito principal analizar los resultados de un determinado experimento aleatorio con el fin de poder predecir futuros resultados si dicho experimento se repite muchas veces.

Sólo que ahora nos referiremos a Modelos correspondientes a Variables Aleatorias Continuas, a quienes asociábamos una Función de Densidad y una Función de Distribución o de Acumulación.

Los tiempos entre llegadas de clientes a una oficina de atención de público, pueden representarse por un modelo *Exponencial*.

El error respecto de una medida especificada, en un objeto producido por un proceso industrial, es una variable continua que puede representarse mediante un modelo *Normal*. También podría representar la dispersión, en torno a un valor promedio, de los puntajes de la prueba de selección universitaria.

Como puede observarse en los casos planteados hay diferentes modelos, todos ellos, aplicables en diferentes situaciones, según distinguimos seguidamente:

Tipo de Variable	***Modelo a seleccionar***
Continuas	*Modelo Uniforme Continuo*
	Modelo Exponencial
	Modelo Normal

1.1. MODELO UNIFORME CONTINUO

Una variable aleatoria cuyo valor sólo puede encontrarse en cierto intervalo infinito (a, *b)*, por ejemplo: x = (*Nro. Real x: a* $\leq x \leq$ *b),* tiene una distribución uniforme, o rectangular, si su función de densidad es constante en dicho intervalo.

Función De Densidad

Entonces, se trata de deducir la función de densidad uniforme en el intervalo (*a, b*), a partir de una constante arbitraria $y = k$.

Siguiendo el procedimiento general para determinar si es función de densidad:

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

1) Integramos entre los límites del intervalo

$$\int_a^b k\,dx = k\,x\Big|_a^b = k(b-a)$$

2) Para que el resultado de la integral dé como resultado 1, dividimos ambos miembros por $k\,(b - a)$.

$$\int_a^b \frac{k}{k(b-a)}\,dx = \frac{k(b-a)}{k(b-a)}$$

O sea, $\int_a^b \frac{1}{b-a}\,dx = 1$, lo cual indica que la función constante $f(x) = \frac{1}{b-a}$ es una *Función De Densidad* constante o uniforme en el intervalo $(a\,,\,b)$ y lo expresamos:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \forall \text{otro } x \end{cases} \qquad (1)$$

De manera que si queremos encontrar la probabilidad de que la variable aleatoria asuma un valor cualquiera en el intervalo $(c\,,\,d)$, siendo $a < c < d < b$ (un subintervalo) será:

$$P(c<x<d) = \int_c^d \frac{1}{b-a}\,dx = \frac{x}{b-a}\Big|_c^d = \frac{d}{b-a} - \frac{c}{b-a} = \frac{d-c}{b-a} = \frac{1}{b-a}(d-c)$$

El coeficiente $\frac{1}{b-a}$, o sea la función de densidad uniforme se suele llamar factor de proporcionalidad y nos indica que la probabilidad es proporcional a la longitud del intervalo cuya probabilidad buscamos.

Entonces, la probabilidad de cualquier hecho que es un subintervalo de (a, b), es precisamente la razón de la longitud de dicho intervalo (d, c) a la longitud de (a, b); y todo evento que no pertenezca al intervalo (a, b) será igual al evento imposible, es decir tendrá probabilidad 0.

Función de Distribución o Función de Acumulación

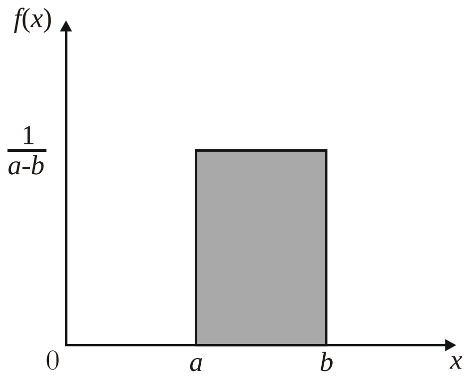
$$F(x) = \int_a^x f(x)\,dx = \int_a^x \frac{1}{b-a}\,dx = \frac{x-a}{b-a} \qquad \text{y la}$$

Y su expresión correcta es:

$$F(x) = \begin{cases} \dfrac{x-a}{b-a} & a < x < b \\ 0 & x \leq a \\ 1 & x \geq b \end{cases} \qquad (2)$$

De (1) y (2) se aprecia que los parámetros de la distribución uniforme serán los valores de a y b.

Gráficamente:



Función De Densidad De Una Distribución Uniforme

Parámetros

Esperanza

Sabemos que $E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$ es la esperanza matemática para variables continuas.

Entonces, reemplazando y resolviendo

$$E(X) = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a} \int_a^b x\, dx = \frac{1}{b-a} \frac{x^2}{2} \Big|_a^b = \frac{1}{(b-a)} \left(\frac{b^2}{2} - \frac{a^2}{2} \right) =$$

$$= \frac{1}{b-a} \frac{b^2 - a^2}{2} = \frac{1}{b-a} \frac{(b-a)(b+a)}{2} = \frac{\cancel{(b-a)}(b+a)}{2\cancel{(b-a)}} = \frac{(b+a)}{2}$$

Luego $$\boxed{E(X) = \frac{a+b}{2}}$$

Varianza

Recordemos que $V(X) = E(X^2) - [E(X)]^2$, por lo tanto calculamos $E(X^2)$; pues $E(X)$ fue calculada anteriormente.

$$E\left(X^2\right) = \int_a^b x^2 \frac{1}{b-a}\, dx = \frac{1}{b-a} \int_a^b x^2\, dx = \frac{1}{b-a} \left. \frac{x^2}{3} \right|_a^b = \frac{1}{b-a}\left(\frac{b^3}{3} - \frac{a^3}{3}\right)$$

$$= \frac{1}{b-a}\frac{\left(b^3 - a^3\right)}{3} = \frac{\left(b^3 - a^3\right)}{3(b-a)}$$

Entonces:

$$V(X) = E\left(X^2\right) - [E(X)]^2 = \frac{\left(b^3 - a^3\right)}{3(b-a)} - \left[\frac{b+a}{2}\right]^2 = \frac{\left(b^3 - a^3\right)}{3(b-a)} - \frac{(b+a)^2}{4} =$$

$$= \frac{4\left(b^3 - a^3\right) - 3(b-a)(b+a)^2}{12(b-a)} = \frac{4b^3 - 4a^3 - 3(b-a)(b+a)(b+a)}{12(b-a)}$$

$$= \frac{4b^3 - 4a^3 - 3b^3 - 3b^2 a + 3a^2 b + 3a^3}{12(b-a)} = \frac{b^3 - a^3 - 3b^2 a + 3a^2 b}{12(b-a)} =$$

$$= \frac{(b-a)^2 (b-a)}{12(b-a)} = \frac{(b-a)^2}{12}$$

Por lo tanto:

$$\boxed{V(X) = \frac{(b \quad a)^2}{12}}$$

Desviación Estándar

$$\boxed{DS(X) = \sqrt{\frac{(b-a)^2}{12}}}$$

Ejemplos:

1) Si queremos definir una función de densidad uniforme, en el intervalo (2,6), tendremos:

$$f(x) = \begin{cases} \dfrac{1}{6-2} = \dfrac{1}{4} & 2 \le x \le 6 \\ 0 & \forall \text{ otro } x \end{cases}$$

$$F(x)=\left|\begin{array}{ll} \dfrac{x-2}{6-2}=\dfrac{x}{4}-\dfrac{1}{2} & 2<x<6 \\ 0 & x\le 2 \\ 1 & x\ge 6 \end{array}\right.$$

$$E(X)=\frac{2+6}{2}=4 \quad \text{y} \quad V(X)=\frac{(6-2)^2}{12}=1{,}33$$

2) Supongamos que contamos con un péndulo que oscila en un intervalo (0,1), es decir este aparato producirá, al detenerse un número aleatorio comprendido entre los límites 0 y 1, de manera que la probabilidad de que el péndulo se ubique dentro de cierto intervalo dado, es proporcional a la longitud de dicho intervalo. Los eventos elementales de este espacio probabilístico serán los números reales entre 0 y 1 y la familia *F* de eventos será aquella que incluya los subintervalos de (0,1) que satisfagan los axiomas para *F.* Entonces, la probabilidad de un evento que corresponde a un subintervalo de (0,1) es proporcional a la longitud de dicho subintervalo. Dado que la probabilidad es igual a 1 cuando la longitud del intervalo es 1, entonces la proporcionalidad constante debe ser igual a 1.

Por otra parte todo evento que no pertenezca al intervalo (0, 1) tiene probabilidad 0, es decir entonces que:

Si $0 < a < b < 1$, será

$$\Pr(a<X<b)=\int_a^b 1\,dx=b-a$$

De manera que la constante 1 es la Función de Densidad de la variable aleatoria *X*, para $0 < X < 1$ es decir

$$f(x)\left|\begin{array}{ll} 1 & \text{si} \quad 0<x<1 \\ 0 & \text{si} \quad x\le 0 \text{ ó } x\ge 1 \end{array}\right. \qquad \int_0^1 1\,dx = x\,\Big|_0^1 = 1$$

Y la Función de Distribución será:

$$F(x)\left|\begin{array}{lll} 0 & \text{para} & x\le 0 \\ x & \text{para} & 0<x<1 \\ 1 & \text{para} & x\ge 1 \end{array}\right. \qquad F(X)=\int_0^1 1\,dx = x\,\Big|_0^x = x-0=x$$

1.2. MODELO EXPONENCIAL

El Modelo De Probabilidad Exponencial tiene su origen en el Modelo De Poisson, donde según vimos una posibilidad se relaciona con la probabilidad de ocurrencia de un número determinado de éxitos en una unidad especificada finita, donde el número de éxitos pasa a ser la variable aleatoria.

Ahora, si invertimos los papeles de una variable de Poisson y su unidad especificada finita, tenemos lo que se llama un Modelo De Probabilidad Exponencial.

Precisamente, una variable exponencial x es el intervalo de tiempo, o espacio, requerido para obtener un número específico de éxitos y ya que el tiempo o el espacio son continuos, una medición de este tipo es una variable aleatoria continua.

Por ejemplo, si las llegadas de automóviles a una casilla de peaje siguen la ley de Poisson, entonces el tiempo transcurrido entre llegadas sucesivas de automóviles es lo que llamaríamos una variable exponencial.

Este modelo es de gran importancia en la teoría de probabilidades aplicada, porque puede usarse para describir una gran variedad de situaciones aleatorias. En las ciencias de la administración se encuentran a menudo problemas de operación tales como determinar el número de casillas de peaje requeridas en las entradas o salidas de carreteras, el número de líneas telefónicas en una zona determinada, el número de trabajadores en una cafetería, etc. Cada una de estas situaciones tiene una característica común, la existencia de una demanda variable de servicio con el tiempo que puede ser satisfecha por un número específico de trabajadores. Esta demanda variable en el tiempo obedece a la ley exponencial. También caracteriza bien a fenómenos como la duración de una cinta electrónica, los intervalos de tiempo entre descomposturas de mecanismos eléctricos o entre accidentes, etc.

De los ejemplos anteriores, vemos que el modelo de probabilidad exponencial surge como respuesta a la pregunta: si una serie de hechos ocurre en tiempo según la Ley De Poisson a un ritmo de λ hechos por unidad de tiempo, ¿cuánto tenemos que esperar para poder observar la primera ocurrencia de un hecho? La respuesta a esta pregunta sugiere un método de elaborar el Modelo De Probabilidad Exponencial partiendo de un Proceso De Poisson. Tal como resulta ser cierto para cualquier variable aleatoria continua, no tiene sentido preguntar "¿Cuál es la probabilidad de que el primer hecho llegue exactamente en un minuto?". Más bien, se debe determinar un intervalo dentro del cual debe ocurrir el hecho, como por ejemplo, preguntado "¿Cuál es la probabilidad de que el primer hecho llegue en un minuto (dentro de)?".

Como los Procesos Poisson son estacionarios, y se tiene una probabilidad igual de que el evento ocurra a todo lo largo del período relevante de tiempo, la Distribución Exponencial se aplica si lo que interesa es el tiempo (o espacio) hasta la ocurrencia del primer hecho, o el tiempo entre dos hechos sucesivos o el tiempo que transcurre hasta que se presenta el primer hecho, después de cualquier punto en el tiempo elegido al azar.

Designaremos por x la variable tiempo entre hechos y procederemos a determinar la función de distribución de x evaluando el hecho $x > x_i$ para cualquier intervalo de tiempo específico x.

El hecho $x > x_i$ no ha ocurrido aún en el intervalo de tiempo (0; x). Por la Ley De Poisson, su probabilidad es:

P($x > x_i$) = P [no ocurrencia en (0; x)] = $e^{-\lambda x}\left[\frac{(\lambda x)^0}{0!}\right] = e^{-\lambda x}$

De esta manera la Función De Acumulación de la variable exponencial x puede expresarse:

$$T(x) = P(x \leq x_i) = 1 - e^{-\lambda x}$$

Puede demostrarse que la Función De Densidad de una variable exponencial es:

$$t(x) = \lambda \; e^{-\lambda x} \quad \text{para } x \geq 0$$

También puede demostrarse que una variable exponencial, como una variable Poisson, tiene un solo parámetro, su media. Para una Distribución Exponencial tenemos:

$$E(x) = \beta = \frac{1}{\lambda}$$

$$V(x) = \beta^2 = \frac{1}{\lambda^2}$$

$$DS(x) = \beta = \frac{1}{\lambda}$$

De la relación entre λ y β, vemos que las Funciones De Densidad y Acumulación de una Variable Exponencial también pueden escribirse, respectivamente, como

$$t(x) = \frac{1}{\beta} \; e^{-\frac{x}{\beta}} \qquad \text{para } x \geq 0$$

y

$$T(x) = 1 - e^{-\frac{x}{\beta}} \qquad \text{para } x \geq 0$$

En las expresiones anteriores, *x* es el intervalo de tiempo entre ocurrencias de la variable Poisson, *e* es la base de los logaritmos naturales, y λ es el promedio de ocurrencias por unidad de tiempo de la variable de Poisson. El hecho de que $\beta = \frac{1}{\lambda}$ es un resultado natural, porque la variable exponencial es el tiempo entre sucesivas ocurrencias de Poisson. A este respecto, vemos que puede ser interpretada como el intervalo de tiempo medio entre ocurrencias de Poisson, o el tiempo esperado hasta la primera ocurrencia de un hecho Poisson. Por otro lado, en la Distribución de Poisson la Esperanza es igual a la Varianza, mientras que en la Distribución Exponencial, la igualdad es entre la Esperanza y la Desviación Estándar.

Este modelo es también llamado exponencial negativo, porque la pendiente de la curva de densidad es siempre negativa, indicando que la probabilidad de largos intervalos de tiempo entre ocurrencias debe ser menor que la probabilidad de intervalos más cortos de tiempo.

Ejemplo 1:

La distribución de vida durante la cual cierta marca de computadora funciona eficazmente (horas de operación eficaz antes de la primera descompostura), es exponencial con $\beta = 360$ horas. ¿Cuál es la probabilidad de que la computadora funcione eficazmente menos de 180 horas? ¿Más de 720 horas?
Resolución:

$$t(x) = \frac{1}{360} e^{-\frac{x}{360}} \qquad (0 \leq x \leq \infty)$$

$$T(x) = 1 - e^{-\frac{x}{360}}$$

Entonces:

$$P(x \leq 180) = T(180) = 1 - e^{-\frac{180}{360}} = 1 - e^{-0,5} = 1 - 0,60653 = 0,39347$$

$$P(x > 720) = e^{-\frac{720}{360}} = 0,13534$$

$$P(180 \leq x \leq 720) = T(720) - T(180) = (1 - 0,13534) - 0,39347 = 0,47119$$

En resumen, como los procesos Poisson son estacionarios, y se tiene una probabilidad igual de que el evento ocurra a todo lo largo del período relevante de tiempo, la distribución exponencial se aplica si lo que interesa es el tiempo (o espacio) hasta la ocurrencia del primer evento, o el tiempo entre dos eventos sucesivos o el tiempo que transcurre hasta que se presenta el primer evento, después de cualquier punto en el tiempo elegido al azar.

La probabilidad exponencial de que ocurra el primer evento dentro del intervalo designado de tiempo es (en donde λ es el número promedio de ocurrencias para el intervalo de interés).

$$\Pr(x \leq x_i) = 1 - e^{-\lambda x}$$

De manera similar, la probabilidad exponencial de que le primer evento no ocurra dentro del intervalo designado de tiempo o espacio es:

$$\Pr(x > x_i) = e^{-\lambda x}$$

Ejemplo 2:

En un departamento de reparación de maquinaria se reciben 5 solicitudes por hora en promedio. Comenzando la observación en cualquier punto del tiempo, la probabilidad que se reciba la primera solicitud de servicio dentro de un lapso de media hora es:

Promedio por hora = 5
λ = Promedio por media hora = 2,5
$P = 1 - e^{-\lambda} = 1 - e^{-2,5} = 1 - 0,08208 = 0,91792$

1.3. MODELO NORMAL

Su importancia se debe principalmente a tres consideraciones generales:

1) Hay muchas variables que parecen seguir una forma de variación que es análoga a la distribución normal.

Muchos aspectos naturales, tales como los pesos y las alturas de los humanos o coeficiente de inteligencia, se consideran como de distribución normal.

Normalmente, esto no significa que una distribución de frecuencia relativa, realmente observada, siga exactamente la distribución normal, pero en promedio, las probabilidades asociadas a la distribución teórica se aplican muy aproximadamente.

La distribución normal es simplemente una teoría que sirve para explicar, para algunas variables aleatorias, la relación entre intervalos de sus valores y sus correspondientes probabilidades.

2) La distribución muestral de muchos estadígrafos muestrales, tales como la media, tiene una distribución aproximadamente normal independientemente de la distribución de la población, si el tamaño de la muestra es lo suficientemente grande.

3) Es una excelente aproximación de otras varias distribuciones muestrales. Así, la distribución Binomial, Poisson, Hipergeométrica, se aproximan a la normal al incrementar el tamaño muestral.

Este modelo puede expresarse en la ***Forma General y en Forma Estandarizada***.

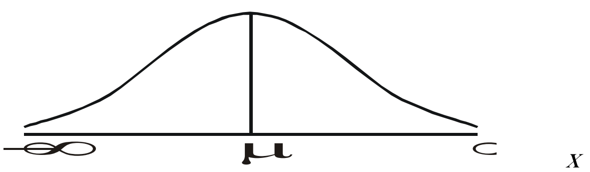
1.3.1. MODELO NORMAL GENERAL

Dada una variable aleatoria *x*, diremos que tiene distribución aproximadamente Normal, con media μ y desviación σ, es decir $x \sim N(\mu,\sigma)$, si su función de densidad es:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad -\infty < x < \infty$$

Donde $e \cong 2{,}71828$ y $\pi = 3{,}1416$ son constantes.

Gráfica:



La curva lisa y llana que se observa en la figura, se llama *Curva Normal.*

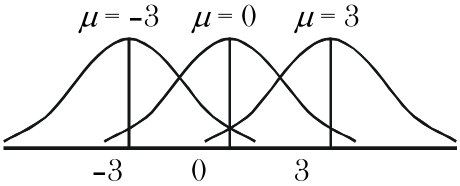
Observemos:

1) El símbolo $f(x)$, representa la densidad de probabilidad en cada posible valor de x, y gráficamente es la ordenada de la curva en todo punto posible x de las abscisas.
2) Toda el área bajo la curva es igual a 1.

3) El área determinada por un intervalo debajo de la curva y por encima del eje de las abscisas, es la probabilidad de ese intervalo.

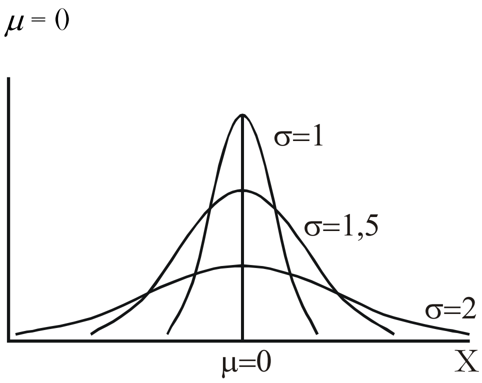
4) La expresión de una función muestra que la densidad de la probabilidad $f(x)$ depende, fuera de x, valor de la variable aleatoria, solamente de dos parámetros, μ y σ^2. El parámetro μ es la media de la distribución y σ^2 su varianza.
   Los valores de μ y σ determinan la situación y forma de la curva normal.
   Dada σ, una variación de μ traslada la curva como un todo a lo largo de un eje x.

Ejemplo:
$\sigma = 1{,}5$



Para un mismo valor central, cuanto mayor es la desviación típica más aplanada es la curva, cuanto menor es la desviación, más acusada es la cresta.

Ejemplo:



Es claro que la distribución está completamente definida por la media μ y la desviación σ. Por ello entonces la distribución que tiene media μ y desviación σ se denota por $x \sim N(\mu, \sigma)$ ó simplemente con parámetros μ y σ, entendiéndose que en el primer parámetro se menciona a la media y en el segundo a la desviación.

5) Teóricamente hablando, el dominio de $f(x)$ es infinito, y es una distribución para todos los valores de x entre $-\infty$ y $+\infty$ y todo intervalo no nulo a lo largo del eje x tiene una probabilidad no nula.

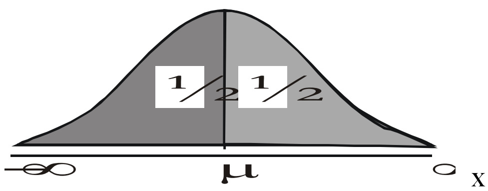
6) La curva va como aproximándose, pero sin tocar nunca el eje x, en medida que x se aleja de μ, por ello se llama curva asintótica.

7) Como la probabilidad de que x tome algún valor exacto x_i es 0, toda expresión de probabilidad que se relacione con una variable tiene que ponerse en forma de intervalo o acumulativa.
Entonces, la probabilidad de que x quede por debajo de un cierto valor de x, es la probabilidad del intervalo entre $-\infty$ y x, o sea:
$F(x) = P(x \leq x_i) = \int_{-\infty}^{x_i} f(x)\,dx$, y constituye la función de probabilidad acumulada de x o función de acumulación.

Propiedades

1- Es *simétrica* con respecto al valor medio, eso implica que las superficies son las mismas desde $-\infty$ a μ que desde μ a ∞, es decir:

$$\int_{-\infty}^{\mu} f(x)\,dx = \int_{\mu}^{\infty} f(x)\,dx = \frac{1}{2}$$



2- Es *unimodal o de forma acampanada.*

Estos atributos se pueden ver fácilmente en la ecuación

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Veamos:

Dada una distribución con σ^2 *conocida*, se sabe que $\frac{1}{\sigma\sqrt{2\pi}}$ es una constante.

El otro factor puede escribirse $\dfrac{1}{e^{\frac{(x-\mu)^2}{2\sigma^2}}}$, donde e, σ^2 y μ son constantes.

Entonces nos interesa el numerador del exponente: $(x-\mu)^2$.

Cuando $x = \mu$, entonces $(x = \mu)^2 = 0$. Como todo número elevado a la potencia 0 es 1, el denominador de la fracción anterior y por lo tanto, la fracción misma, es 1.

El valor de la función en el punto x = μ es $\frac{1}{\sigma\sqrt{2\pi}}$. Este valor es, por lo tanto, la altura de la curva en el punto central $x = \mu$, es el único que da la ordenada máxima.

Cuando x se aleja de μ, la diferencia $x - \mu$ se hace mayor, y por lo tanto también el cuadrado de la diferencia, lo que equivale a que la fracción: $\frac{1}{e^{\frac{(x-\mu)^2}{2\sigma^2}}}$ se haga menor.

En consecuencia, la altura también será menor. Pero la altura menor no se anula sino cuando la desviación de x respecto a $\mu \to \infty$. Obsérvese que la diferencia entre x y μ interviene como un valor al cuadrado. De manera que dos valores de x diferentes que difieran lo mismo de μ, tendrán la misma densidad, cosa que explica la simetría de la curva normal.

La media, la mediana, y la moda son idénticas.

3- Transformación lineal de escala

Si x y y son dos variables aleatorias y $y = a + x$, la varianza de y es la misma y la media de y es simplemente la media de $a + x$.

Es de esperar que si x tiene distribución normal con μ y σ^2, la variable y también sea normal con media $\mu + a$ y varianza σ^2.

En general:

$y = a + b.\,x$ y $x \sim N(\mu, \sigma^2)$, entonces $y \sim N(a + b.\mu;\, b^2.\sigma^2)$.

Es decir que una transformación lineal de escala de una variable x con distribución normal, da una nueva variable y que también tiene distribución normal.

4- Combinación lineal de variables aleatorias

Si x_1, x_2, ... , x_n, son variables normales independientes, su suma S, es también una variable normal. Además debido a la independencia, la propiedad aditiva se verifica para la Esperanza y Varianza. Es decir, la esperanza de S es la suma de las esperanzas de las n variables normales. Igualmente la Varianza.

Esperanza, Varianza y Desviación Estándar

$$E(x) = \int_{-\infty}^{\infty} x\, f(x)\, dx \quad \text{y}$$

$$V(x) = \int_{-\infty}^{\infty} \left[x - E(x)\right]^2 f(x)\, dx$$

$$DS(x) = \sqrt{V(x)}$$

3.2. MODELO NORMAL ESTÁNDAR

También llamada Distribución Normal Típica.

Podemos contar, como hemos visto, con distribuciones con iguales dispersiones pero distintas medias, o bien con iguales medias pero distintas dispersiones.

Esto hace que para cada caso, o para cada estudio particular, habría que determinar su función de densidad y su correspondiente función de distribución o acumulación, o construir infinitas tablas para las infinitas combinaciones de μ y σ.

Para obviar este inconveniente, y a los fines de trabajar con una sola tabla de probabilidades correspondientes a distribuciones normales, cualquiera sean los valores de las medias y las varianzas y mas aún para cualquier unidad de medida, se obtiene la función de densidad normal, para la variable desvío estandarizada, tipificada o reducida, que según vimos se simboliza y define:

$$z = \frac{x-\mu}{\sigma} \sim N(0,1)$$

Y se confecciona la tabla para dicha función. Entonces:

1- La función de densidad queda reducida a la siguiente expresión:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}; \qquad z \approx N(0,1)$$

y tiene las mismas características que la función general.

2- $E(z) = \int_{-\infty}^{\infty} z\, f(z)\, dz$ y

$$V(z) = \int_{-\infty}^{\infty} [z - E(z)]^2 f(z)\, dz = \int_{-\infty}^{\infty} z^2 f(z)\, dz \qquad para\ E(z) = 0$$

Demostraremos a continuación que $E(z) = 0$ y $V(z) = 1$

$$z = \frac{x-\mu}{\sigma} = \frac{1}{\sigma}x - \frac{\mu}{\sigma}$$

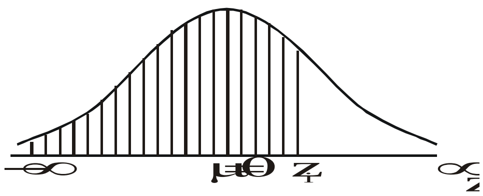
Como $1/\sigma$ y μ/σ son constantes, se tiene:

$$E(z) = E\left(\frac{1}{\sigma}x - \frac{\mu}{\sigma}\right) = \frac{1}{\sigma}E(x) - \frac{\mu}{\sigma} = \frac{1}{\sigma}\mu - \frac{\mu}{\sigma} = 0$$

$$V(z) = V\left(\frac{1}{\sigma}x - \frac{\mu}{\sigma}\right) = \frac{1}{\sigma^2}V(x) - 0 = \frac{1}{\sigma^2}\sigma^2 - 0 = 1$$

Función de Distribución o Función de Acumulación.

Nos proporciona la probabilidad de que la variable aleatoria asuma un valor cualquiera en el intervalo $(-\infty; z)$.



$$F(z) = P\left(z \le z_i\right) = \int_{-\infty}^{z_i} f(z)\,dz \qquad z \sim N(0,1)$$

Si queremos encontrar la distribución de probabilidad de un intervalo dado $[a - b]$, se encuentran los valores de z para a y b y luego se obtiene la probabilidad tomando la diferencia entre las probabilidades acumuladas de los valores de z.

$$P(a < x < b) = P\left(\frac{a-\mu}{\sigma} < z < \frac{b-\mu}{\sigma}\right)$$

Tablas Usuales

La tabla con la que trabajaremos, (TABLA VI) corresponde a una variable z, es decir con media 0 y desviación 1 $[z \sim N(0,1)]$, por ello toda variable $x \sim N(\mu,\sigma)$, deberá ser previamente transformada a una $z \sim N(0,1)$ a los fines de poder utilizar esta tabla.

Además, acumula desde $-\infty$ hasta un valor positivo de la variable, o sea las probabilidades en el cuerpo de la tabla corresponden a:

$$F(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\,dz = Pr\left(z < z_i\right)$$

<u>Ejemplos</u>:

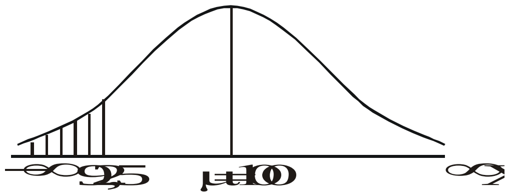
1- Sea x una $N^{\sim}(100,15)$, hallar las probabilidades siguientes:

a- $P(x < 92,5)$
b- $P(x > 76)$
c- $P(x < 107,5)$
d- $P(x > 124)$
e- $P(77,5 < x < 100)$
f- $P(112 < x < 128,5)$

Del anterior enunciado deducimos que $\mu = 100$ y $\sigma = 15$, pues dijimos que el primer valor corresponde a la media y el segundo a la desviación.

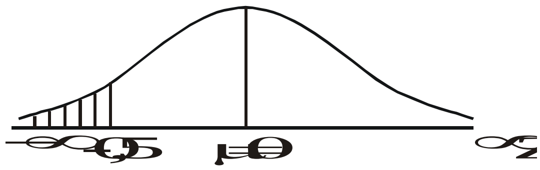
Como nos plantean la variable *x*, debemos primeramente transformarla a *z*, a los fines de utilizar la tabla.

a- *P (x < 92,5)*



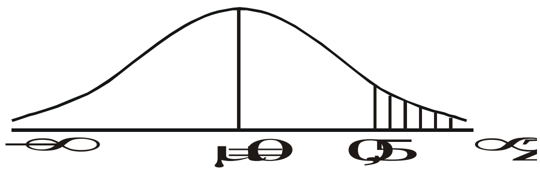
Transformamos a z:

$$P\left(\frac{x-\mu}{\sigma} < \frac{92,5-100}{15}\right) = P(z < -0,5)$$



Como la tabla acumula desde –∞ hasta un valor positivo de la variable, la anterior probabilidad no puede encontrarse directamente, por lo cual teniendo en cuenta la propiedad de simetría, resulta que:

Pr (z < -0,5) = Pr (z > 0,5)



Y el segundo miembro se resuelve razonando así:

Si la probabilidad total (desde –∞ hasta∞) vale 1 y encontramos en tabla:

Pr (z <0,5), restándole esta última al 1, encontramos la probabilidad buscada, o sea:

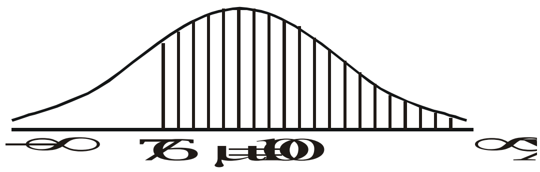
Pr (z < -0,5) = Pr (z > 0,5) = 1 – Pr (Z <0,5), o bien

$\int_{-\infty}^{-0,5} f(z)\,dz = \int_{0,5}^{\infty} f(z)\,dz = 1 - \int_{-\infty}^{0,5} f(z)\,dz$, que es otra manera de expresar las anteriores probabilidades.

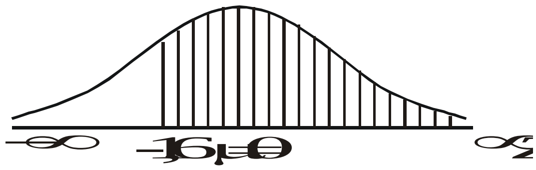
Entonces:

Pr (z < -0,5) = Pr (z > 0,5) = 1 – 0,6915 = <u>0,3085</u>

b- P (x > 76)



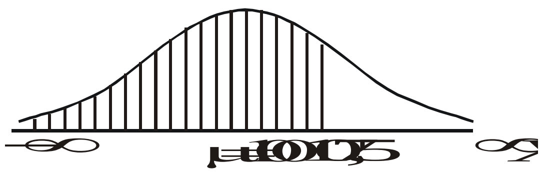
$$\Pr\left(\frac{x-\mu}{\sigma} > \frac{76-100}{15}\right) = \Pr(z > -1{,}6)$$



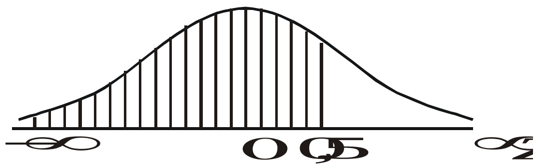
Entonces:

$$\Pr(z > -1{,}6) = \Pr(z < 1{,}6) = \underline{0{,}9452}$$

c- P (x < 107,50)

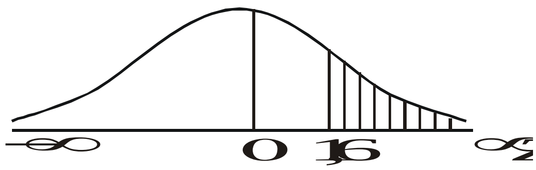


$$\Pr\left(\frac{x-\mu}{\sigma} > \frac{107{,}5-100}{15}\right) = \Pr(z < 0{,}5) = 0{,}6915$$



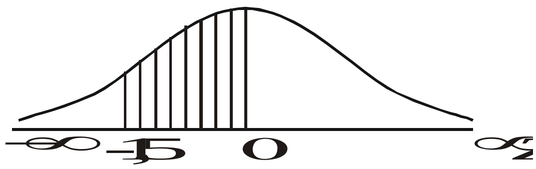
d- P(x > 124)

$$\Pr\left(\frac{x-\mu}{\sigma} > \frac{124-100}{15}\right) = \Pr(z > 1{,}6) = 1 - \Pr(z < 1{,}6) = 1 - 0{,}9452 = 0{,}0548$$



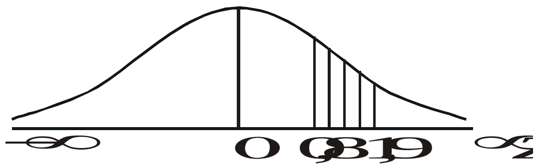
e- P (77,5 < x < 100)

$$\Pr(-1{,}5 < z < 0) = \Pr(z < 1{,}5) - \Pr(z < 0) = 0{,}9332 - 0{,}50 = 0{,}4332$$



f- P (112 $<$ x $<$ 128,5)

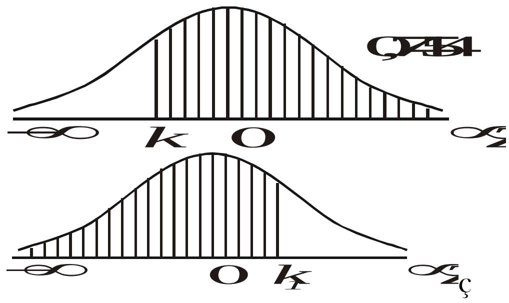
Pr $(0{,}80 < z < 1{,}9)$= Pr $(z < 1{,}9)$ – Pr $(z < 0{,}80)$ = 0,9713 – 0,7881 = 0,1832



2- En el anterior ejercicio se nos pedía el cálculo de probabilidades, pero puede ocurrir que tengamos como dato la probabilidad y necesitemos encontrar el valor de *Z*.

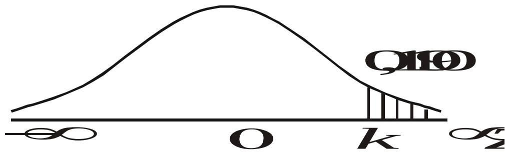
Así:

a- Pr $(z > k) = 0{,}7454$



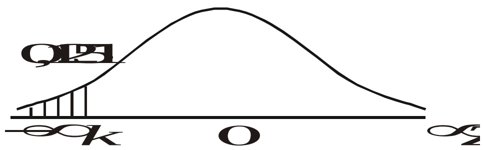
Pr $(z > k)$ = Pr $(z < k_1)$ = 0,7454 $\Rightarrow$ k_1 = 0,66 y k = –0,66

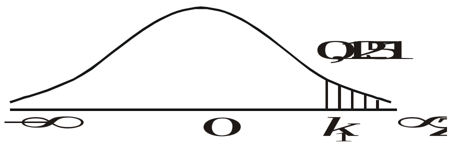
b- Pr $(z > k) = 0{,}1190$



Pr $(z < k)$ = 1– Pr$(z > k)$
0,8810 = 1 – 0,1190
k = 1,18

c- Pr $(z < k) = 0{,}1251$



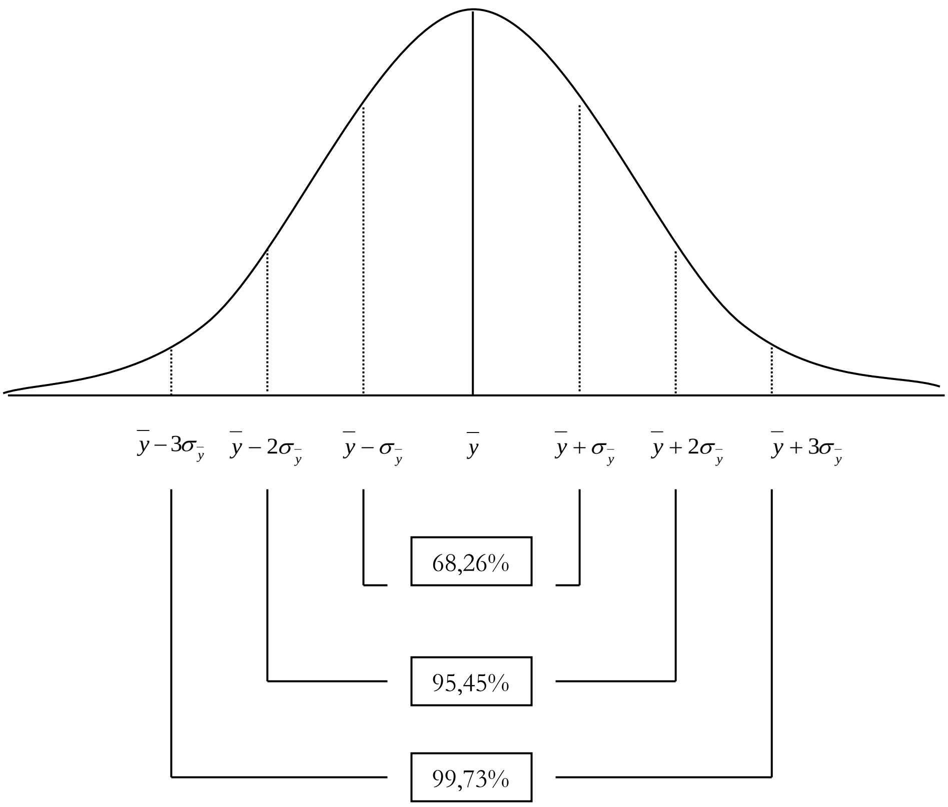


$\Pr(z < k) = \Pr(z > k_1) = 0{,}1251$
$1 - \Pr(z > k_1) = \Pr(z < k_1)$
$1 - 0{,}1251 = 0{,}8749 \Rightarrow k_1 = 1{,}15$
$k = -1{,}15$

1.3.3. APLICACIONES

Regla Empírica Para La Interpretación De La Varianza

Cuando n es grande y la distribución de las n observaciones es aproximadamente en forma de campana, aproximadamente Normal, puede usarse la regla empírica. En este caso las probabilidades están distribuidas alrededor de la media según los porcentajes del gráfico.



Es decir, que si la distribución de frecuencias de un problema analizado tiene el comportamiento similar al de la distribución normal, podemos sacar iguales conclusiones en cuanto a los porcentajes de frecuencias, u observaciones, que se encuentran comprendidos en determinados intervalos de la variable.

Ejemplo:

Supongamos tener un grupo de 110 operarios clasificados por sus jornales horas, con la siguiente distribución:

$y'_{i-1} - y'_i$	y_i	n_i
52,5-57,5	*55*	*5*
57,5-62,5	*60*	*12*
62,5-67,5	*65*	*21*
67,5-72,5	*70*	*25*
72,5-77,5	*75*	*29*
77,5-82,5	*80*	*11*
82,5-87,5	*85*	*4*
87,5-92,5	*90*	*3*
		110

1 — *68%*

2 — *93%*

3 — *100%*

Calculamos:

Media Aritmética	*M (y)* = 69,5
Varianza	*V (y)* = 68,75
Desviación Estándar	*DS (y)* = 8,3

Con estos valores determinamos los intervalos notables:

1°) 69,5 ± 8,3 ⇒ 61,2 ; 77,8 [61,2; 77,8]

2°) 69,5 ± 2(8,3) ⇒ 52,9 ; 86,1 [52,9; 86,1]

3°) 69,5 ± 3(8,3) ⇒ 44,6 ; 94,4 [44,6; 94,4]

Si nos fijamos en la distribución de frecuencias de la tabla, dentro del primer intervalo [61,2; 77,8], determinado y tomado como suficientemente aproximado el intervalo [62,5; 77,5] de la tabla, tenemos comprendido el 68% de las frecuencias (Total: 110. Frecuencias en el intervalo considerado: 21+25+29 =75, entonces 75/110=68%).

De acuerdo a la distribución teórica, se deberían haber encontrado 68 frecuencias, , pero si consideramos en nuestro ejemplo:

1- Que estamos ante una distribución que no es perfectamente simétrica.
2- Que el intervalo considerado [62,5; 77,5] es menor que el calculado [61,2; 77,8], podemos afirmar, sin temor a equivocarnos, que nuestro ejemplo confirma lo que habíamos visto gráficamente.

Para el segundo intervalo [52,9; 86,1] tomamos [52,5; 87,5], que resulta de mayor recorrido que el anterior. Por eso, consideramos tres de las cinco frecuencias en el primer intervalo y sólo dos de las cuatro del penúltimo, con el cual obtenemos un 93% (103/110), cifra que es inferior al 95,45% teórico.

En el tercer intervalo [44,6; 94,4] se presentan valores que se encuentran por debajo y por arriba de los valores de la tabla [52,5; 92,5], por lo que afirmamos que en dicho recorrido se encuentra el 100% de la frecuencia, que es casi coincidente con el 99,73%.

Puntaje Típico

Si una distribución de sueldos $y_{(1)_i}$ tiene como media $\overline{y}_{(1)} = 1.000$ y como $DS(y)_{(1)} = 500$, una observación cualquiera $y_{(1)_i}$ = 1.500 dará $z_1 = \frac{1.500 - 1.000}{500} = 1$.

Si otra distribución de pesos, de aves, por ejemplo, tiene $\overline{y}_{(2)} = 2{,}100\, kgs.$ y como $DS(y)_{(2)} = 0{,}800\, kgs.$, una observación cualquiera $y_{(2)_i}$ = 2,900 *kgs.* dará $z_2 = \frac{2{,}900 - 2{,}100}{0{,}800} = 1$.

En ambos casos, el valor desvío estandarizado es igual a 1, por lo tanto, dos observaciones tendrán la misma probabilidad de presentación dentro de sus distribuciones respectivas.

<u>Ejemplo</u>

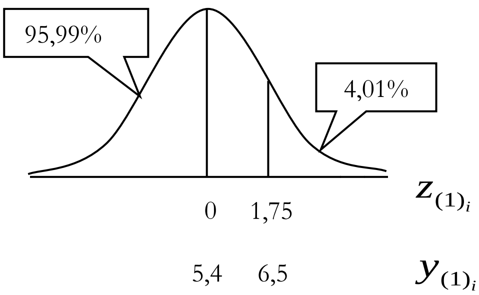
Notas de dos estudiantes en el curso A y en el curso B.

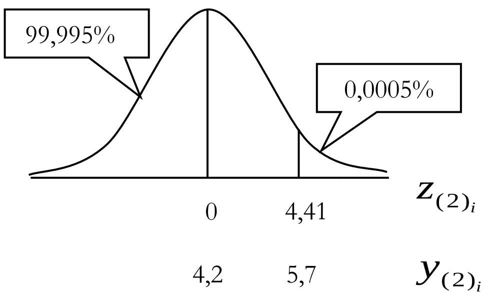
Curso A	*Curso B*
$\overline{y}_{(1)} = 5{,}4$	$\overline{y}_{(2)} = 4{,}2$
$DS(y)_{(1)} = 0{,}63$	$DS(y)_{(2)} = 0{,}34$
$y_{(1)_i}$ = 6,5	$y_{(2)_i}$ = 5,7

Comparadas las notas aisladamente, arribamos a la conclusión de que el alumno A, es mejor que el alumno B. (6,5 > 5,7)

Pero veamos qué ocurre utilizando el Puntaje Típico:

$$z_1 = \frac{6{,}5 - 5{,}4}{0{,}63} = 1{,}75 \qquad\qquad z_2 = \frac{5{,}7 - 4{,}2}{0{,}34} = 4{,}41.$$





La *DS (A)* = 0,63 es más grande que la *DS (B)* = 0,34, lo que implica que las notas del curso A están más dispersas que las notas del curso B.

Tenemos dos valores centrales en ambas distribuciones: uno es el valor de la distribución obtenido en base a las observaciones de las notas del curso, y el otro es el resultado de haber estandarizado la variable desvío.

En el eje horizontal tenemos los valores de z_i, que siguen siendo valores de la variable, pero ahora transformada de escala, puesto que hemos tomado las desviaciones respecto a la media, dividida a su vez, por la desviación estándar.

La zona del primer gráfico, correspondiente a todos los alumnos que han tenido una nota de 6,5 o inferior, y que representan (desde -∞ hasta 1,75) el 95,99% de los casos.

La diferencia, 4,01% representa el porcentaje de los alumnos que han sacado notas superiores al alumno que estamos considerando.

Quiere decir, que el alumno del curso A, tiene por encima un 4,01% de alumnos con nota superior a la de él.

En cambio, en el curso B, las frecuencias a la izquierda del puntaje típico (4,41) es del 99,9995% del total. Quiere decir, que el número de alumnos que podrían tener notas superiores a la que estamos considerando, del grupo B, es insignificante, lo que permite afirmar que el alumno al cual nos referimos es el mejor del curso B, porque por arriba de él hay un 5 por diez mil, porcentaje que no tiene significación, puesto que si el curso ha sido de por ejemplo, 100 alumnos, no habría ninguno con nota superior.

La conclusión es inversa a la primera, puesto que el alumno B ha tenido un mayor rendimiento relativo que el alumno A.

1.4. RELACIÓN ENTRE MODELOS DISCRETOS y EL MODELO NORMAL

Al exponer la Distribución Normal (Modelo Continuo), se dijo que es una excelente aproximación de otras varias distribuciones muestrales, tales como la Distribución Binomial, Poisson, e Hipergeométrico, cuando se incrementa el tamaño de la muestra. Tal aproximación se basa en el Teorema Central del Límite que demuestra esta proximidad. A su vez en el desarrollo de cada una de estas distribuciones discretas hemos mencionado algo al respecto. Analizaremos en este punto con mayor profundidad las relaciones entre las Distribuciones Discretas vistas y la Distribución Normal.

1.4.1. Aproximación del Modelo Binomial para la variable x, al Modelo Normal

Dijimos en el presente capítulo que la forma de la Distribución Binomial, depende de los valores de P y n.

Así:

1- Si $P = Q = 0{,}50$, la distribución será simétrica, independientemente del valor de n.

2- Si $P \neq Q$, la distribución será asimétrica, entonces para una n dada:

2.1. Si $P < Q$ será asimétrica positiva o derecha

2.2. Si $P > Q$ será asimétrica negativa o izquierda.

Sin embargo, al incrementar el tamaño de la muestra (n), la distribución se hace menos asimétrica, aproximándose a la simetría independientemente de la diferencia entre P y Q cuando el tamaño muestral tiende a infinito ($n \to \infty$).

Así, a medida que el valor de n aumenta, la *Distribución Binomial* se convierte en una curva continua y simétrica sean o no iguales P y Q. Tal distribución continua, simétrica y acampanada es la *Distribución Normal*, que es el límite de la Binomial, cuando n tiende a infinito.

Entonces, si $x \sim \beta(n, P)$, con $E(x) = n.P$

$V(x) = n.P.Q \Rightarrow \sigma_x = \sqrt{nPQ}$

Como aproximación a la Normal cuando $n \to \infty$, habrá que definir $z = \dfrac{x - E(x)}{\sigma_x}$,

reemplazando a $E(x)$ y σ_x por su igual según la distribución Binomial, luego

$$z = \frac{x - nP}{\sqrt{nPQ}} \sim N(0,1).$$

Debe tenerse en cuenta que en la transformación de una variable Binomial, que es discreta, en normal tipificada, que es continua, es necesario hacer una corrección dc continuidad.

En general, si x es una variable Binomial cuyas probabilidades se desean conocer, se pueden aproximar las probabilidades aplicando la siguiente corrección de continuidad:

$$P\,(a \leq x \leq b) \approx P\,(a - 0{,}5 \leq x' \leq b + 0{,}5)$$

Donde x' es la variable normal transformada.

Teniendo en cuenta esta corrección, procederemos ahora a demostrar, a través de un ejemplo, que la aproximación normal de la binomial se hace mejor a medida que el tamaño de la muestra se hace mayor.

Supongamos que $x \sim \beta\,(10;\,0{,}5)$ y que se desea hallar la $P\,(2 \leq x \leq 4)$ ó bien $\sum_{x=2}^{4} P(x)$

Resolveremos:

1) Aplicando la Distribución Binomial.

Siendo $n = 10$ y $P = 0{,}5$; entonces:

Pr $(2 \leq x \leq 4 \;\; n = 10 \;\; P = 0{,}5)$ = =Pr $(x \leq 4 \;\; n = 10 \;\; P = 0{,}5)$ – Pr $(x \leq 1 \;\; n = 10 \;\; P = 0{,}5)$ = =<u>0,3663</u>

2) Aplicando la distribución normal

Es necesario primero, la corrección de continuidad:

$$P\,(2 \leq x \leq 4) \approx P\,(2 - 0{,}5 \leq x' \leq 4 + 0{,}5)$$

Entonces buscamos $P\,(1{,}5 \leq x' \leq 4{,}5)$, para lo cual debemos transformar a x' en z.

Siendo $n = 10$ y P = 0,5 $\Rightarrow$ $\quad n.P = 10 \times 0{,}5 = 5$

$$n\,.P.Q = 10 \times 0{,}5 \times 0{,}5 = 2{,}5$$

Luego $$z = \frac{x' - 5}{\sqrt{2{,}5}}$$

Y buscamos $$\Pr\left(\frac{1{,}5-5}{\sqrt{2{,}5}} \leq z \leq \frac{4{,}5-5}{\sqrt{2{,}5}}\right) = \Pr(-2{,}215 \leq z \leq -0{,}315) =$$

$$= \Pr(z < -0{,}315) - \Pr(z < -2{,}215)$$

$$= \left[1 - \Pr(z < 0{,}315)\right] - \left[1 - \Pr(z < 2{,}215)\right]$$

$$= (1 - 0{,}6236) - (1 - 0{,}9866) = 0{,}3630$$

Si comparamos con el resultado exacto (0,3662) la diferencia es de 0,0032.

Supongamos ahora, que en el ejemplo anterior, P permanece constante pero se incrementa a 30 el tamaño muestral (n = 30).

Aplicamos la distribución binomial. La variable binomial es $X \sim \beta$ (30; *0,5)* y se pide:

$$\Pr(2 \le x \le 4,\ n = 30,\ P = 0{,}50) =$$

$$= C_{30}^{2}(0{,}5)^{2}(0{,}5)^{28} + C_{30}^{3}(0{,}5)^{3}(0{,}5)^{27} + C_{30}^{4}(0{,}5)^{4}(0{,}5)^{26} = 0$$

Aplicamos la distribución normal.

$$\Pr(2 \le x \le 4) \cong \Pr(1{,}5 \le x' \le 4{,}5)$$

Además: $nP = 30 \times 0{,}5 = 15$ y $nPQ = 30 \times 0{,}5 \times 0{,}5 = 7{,}5$

Entonces:

$$\Pr\left(\frac{1{,}5-15}{\sqrt{7{,}5}} \le z \le \frac{4{,}5-15}{\sqrt{7{,}5}}\right) = \Pr(-4{,}93 \le z \le -3{,}83) = \Pr(3{,}83 \le Z \le 4{,}93) =$$

$$= \Pr(z \le 4{,}93) - \Pr(z \le 3{,}83) = 1 - 1 = 0$$

Si comparamos con el resultado exacto, no existe diferencia. El valor de n para una aproximación normal satisfactoria dependerá dcl valor de *P*, así:

Si, $P > 0{,}5$, entonces $n(1 - P) > 5$ ó si $P \le 0{,}5$ entonces $nP > 5$, según lo cual para $P = 0{,}5$, *n* debe ser mayor que 10 y para $P = 0{,}9$ ó 0,1, n ha de ser mayor que 50. En general, cuanto más acentuada es la simetría, mayor ha de ser el tamaño de la muestra.

Para simplificar trabajamos con la Normal como aproximación de la Binomial cuando n sea superior a 30 y omitiremos la corrección por continuidad.

Ejemplo:

En cierto proceso de producción el 10% de las unidades producidas son defectuosas. Sea *x* el número de unidades defectuosas halladas en una muestra de 100 unidades tomadas al azar. Hallar los valores aproximados de:

a) $P(x \ge 13)$ c) $P(7 \le x \le 16)$
b) $P(x \ge 8)$ d) $P(5 \le x \le 10)$

Resolución:

$x \sim \beta(n, P)$ Siendo $n = 100$ y $P = 0{,}10$

Debemos transformarla en Normal.

$$x \sim E\left(\mu = nP;\ \sigma = \sqrt{nPQ}\right) \quad \text{Siendo} \quad \begin{array}{l} nP = 100 \times 0{,}10 = 10 \\ \sqrt{nPQ} = \sqrt{100 \times 0{,}10 \times 0{,}90} = 3 \end{array}$$

Entonces:

a) $$\Pr\left(z \geq \frac{13-10}{3}\right) = \Pr(z > 1) = 1 - \Pr(z < 1) = 1 - 0{,}8413 = 0{,}1587$$

b) $$\Pr\left(z \geq \frac{8-10}{3}\right) = \Pr(z > -0{,}67) = \Pr(z < 0{,}67) = 0{,}7486$$

c) $$\Pr\left(\frac{7-10}{3} \leq z \leq \frac{16-10}{3}\right) = \Pr(-1 \leq z \leq 2) = \Pr(z < 2) - [1 - \Pr(z < 1)] =$$
$$= 0{,}9772 - (1 - 0{,}8413) = 0{,}8185$$

d)

$$\Pr\left(\frac{5-10}{3} \leq z \leq \frac{10-10}{3}\right) = \Pr(-1{,}67 \leq z \leq 0) =$$
$$= \Pr(z < 0) - [1 - \Pr(z < 1{,}67)] =$$
$$= 0{,}5 - (1 - 0{,}9525) = 0{,}4525$$

1.4.3. *Aproximación del Modelo Hipergeométrico para la variable x al Modelo Normal*

Si $x \sim H(N, n, X)$

$$E(x) = n\frac{X}{N} \quad \text{y} \quad \sigma_x = \sqrt{\frac{N-n}{N-1} n \frac{X}{N}\left(1 - \frac{X}{N}\right)}$$

$$Z = \frac{x - n\frac{X}{N}}{\sqrt{\frac{N-n}{N-1} n \frac{X}{N}\left(1 - \frac{X}{N}\right)}} \sim N(0,1)$$

1.4.4. *Aproximación del Modelo Binomial e Hipergeométrico para la Proporción Muestral, al Modelo Normal*

La distribución de muestreo de $p(\hat{P})$ es aproximadamente normal si el tamaño de la muestra es lo suficientemente grande; si esta muestra ha sido tomada con reposición.

Para una muestra al azar, tomada sin reposición, de una población finita, el tamaño de la población debe ser considerablemente mayor que el tamaño de la muestra para que esto sea aplicable, sin embargo, este será, por lo general, el caso en la práctica. (En este caso la distribución de $\hat{P}$ se aproxima al modelo Binomial, el cual a su vez, se aproxima al modelo normal como limite).

Al trabajar con variables discretas y aplicarles la distribución Normal, que se refiere a variables continuas, debemos, para obtener aproximaciones normales, satisfactorias, introducir un factor de corrección de continuidad dc 1/2 n (Se emplea 1/2 n en lugar de ½ porque una proporción de éxitos es el número de éxitos dividido por n).

Con el factor de corrección de continuidad, la función de distribución de $\hat{P}$ con n grande es:

$$P\left(\hat{P} = \hat{P}_0\right) \cong N\left(\frac{\hat{P}_0 + \frac{1}{2n} - P}{\sigma_{\hat{P}_0}}\right)$$

Una regla práctica usada con mucha frecuencia para indicar situaciones en las cuales la distribución de muestreo de $\hat{P}_0$ puede ser aproximada mediante la distribución normal, establece que la aproximación normal es apropiada si tanto n $(1 - P)$ y nP son mayores que 5.

Luego, si la Proporción Muestral proviene de una Binomial, y el tamaño de muestra es igual o mayor a 100, o bien igual o mayor a 30 si P=Q, tenderá a la Normal, y aplicando la corrección por continuidad, nos queda:

Si $\frac{x}{n} \sim \hat{P}$ con $\qquad E\left(\hat{P}\right) = P \qquad$ y $\qquad \sigma_{\hat{P}} = \sqrt{\frac{PQ}{n}}$

Entonces $z = \dfrac{\hat{P} - P}{\sqrt{\frac{PQ}{n}}} \sim N(0,1)$

Mientras, que si la Proporción Muestral proviene de una Hipergeométrica, y el tamaño de muestra es igual o mayor a 100, tenderá a la Normal, y aplicando la corrección por continuidad, nos queda:

$$E\left(\hat{P}\right) = P$$

$$\sigma_{\hat{P}} = \sqrt{\frac{PQ}{n}}\sqrt{\frac{N-n}{N-1}}$$

$$z = \frac{\hat{P} - P}{\sqrt{\frac{PQ}{n}\frac{N-n}{N-1}}} \sim N(0,1)$$

1.4.5. *Aproximación del Modelo Poisson para la variable x al Modelo Normal*

La Distribución Normal también es el límite de la Distribución de Poisson.

Dado el valor de *P*, entonces *nP* aumenta al aumentar *n*. A medida que μ , o *nP* aumenta, la distribución de Poisson se acercará cada vez mas a la curva continua acampanada. La utilizaremos cuando $\lambda > 20$, o bien $\lambda \geq 21$.

Entonces sí $x \sim P(\lambda)$ $\quad E(x) = nP$

$$V(x) = nP \Rightarrow \sigma_x = \sqrt{nP}$$

Luego $z = \dfrac{x - nP}{\sqrt{nP}} \sim N(0,1)$

1.5. DISTRIBUCIONES DE LAS PEQUEÑAS MUESTRAS

En general, la teoría del muestreo se refiere a una clase de situaciones en las que la distribución de probabilidades de una estadística de muestra, es normal o aproximadamente normal, o porque el tamaño de la muestra es suficientemente grande para que el Teorema Central del Límite sea aplicable (n ≥ 30 en la estimación de la media poblacional; n≥ 100 en la estimación de la varianza poblacional y n≥ 30, con P = 0,50 en la proporción poblacional).

Pero, a veces, los supuestos de la teoría del muestreo grande no se cumplen, debido a la presencia de situaciones en la que la desviación estándar de la población no es conocida, pues nos enfrentamos a un nuevo problema o a una nueva teoría; o bien el tamaño de la muestra es pequeño debido a limitaciones físicas como las relacionadas con la investigación médica o con los costos de realizar las observaciones.

Si bien con muestras pequeñas no contamos con la cantidad de información deseada, no carecemos completamente de recursos, y debemos a partir de ellos, extraer conclusiones o tomar decisiones. Los procedimientos de inferencia estadística con muestras pequeñas son iguales que los presentados, con la salvedad de que no será posible la aplicación del Teorema Central del Límite ni la suposición de que las distribuciones por muestreo son normales.

El estudio de inferencias estadísticas con pequeñas muestras se llama teoría del muestreo pequeño o teoría del muestreo exacto.

Entonces, la principal diferencia entre teorías de muestreo grande y pequeño es entre distribuciones por muestreo: para muestras grandes, las distribuciones por muestreo son normales; para muestras pequeñas, la distribuciones por muestreo difieren de un caso a otro.

Además, de las distribuciones Binomial y de Poisson, hay tres distribuciones de probabilidades que a menudo son asumidas por una estadística con n pequeña:

χ^2 *Chi-Cuadrado*
t *t de Student*
F *F de Snedecor*

Los tres modelos se relacionan con el modelo de probabilidad normal y se definen por el número de grados de libertad, concepto que presentaremos seguidamente.

Grados de Libertad

Es el nombre dado al número de observaciones linealmente independientes que ocurren en una suma de cuadrados.

Abordaremos la definición de la siguiente manera:

$$5 = \bar{x} = \frac{1}{2}(x_1 + x_2)$$

Es decir: $x_1 + x_2 = 10$

De tal manera, que podría ser:

$x_1 + x_2 = 8 + 2 = 10$ $\quad$ $x_1 + x_2 = 4 + 6 = 10$ $\quad$ y así sucesivamente.

Lo que se observa, es que una vez fijado el valor de uno de los números, el segundo queda automáticamente determinado, ya que el promedio debe ser 5.

En forma semejante, tendríamos que para n números, $x_1, x_2, \ldots x_n$ y dada la condición de que el promedio sea igual a un cierto valor $\bar{x} = k$, donde k es una constante, solamente contaremos con la posibilidad de elegir libremente n-1 valores, quedando automáticamente establecido el restante. Decimos entonces, que contamos con n-1 grados de libertad y lo representamos por φ=δ=n-1.

Diremos que el número de elementos que pueden ser elegidos libremente o el número de variables independientes, indican los grados de libertad existentes en la función.

Otros autores los definen de acuerdo al número de parámetros a estimar. Es por ello que para el caso $\sum(x_i - \mu)^2$, trabajamos con n grados de libertad, ya que no hay ningún parámetro a estimar.

1.5.1. Distribución Chi-Cuadrado (χ^2)

Fundamentación

Si una variable aleatoria X es *N* (μ, σ), entonces la variable tipificada

$$z = \frac{x-\mu}{\sigma} \approx N(0,1)$$

Si elevamos al cuadrado a Z, su distribución, ¿sigue siendo normal?

$$z^2 = \frac{(x-\mu)^2}{\sigma^2} \approx \chi^2_{(1)}$$

La respuesta es negativa. Tiene distribución χ^2 con 1 grado de libertad, ya que se obtuvo un solo valor de X.

La variable χ^2 tiene una familia de distribuciones, una para cada número de grados de libertad.

Si se hacen independientemente dos observaciones, o sea n=2, se obtiene:

$$z_1^2 = \frac{(x_1-\mu)^2}{\sigma^2} \qquad \text{y} \qquad z_2^2 = \frac{(x_2-\mu)^2}{\sigma^2}$$

Entonces:

$$z_1^2 + z_2^2 = \sum_{i=1}^{2} \frac{(x_i-\mu)^2}{\sigma^2} \approx \chi^2_{(2)}$$

Luego, para el caso general, de n observaciones, las sumas de los cuadrados de las z_n, debe tener n grados de libertad.

$$\sum_{i=1}^{n} z_i^2 = \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{\sigma^2} \approx \chi^2_{(n)}$$

Entonces diremos que: *La suma de los cuadrados de n variables aleatorias muestrales independientes con distribución muestral estandarizada tiene distribución χ^2 con φ=δ=n grados de libertad.*

Hay un número infinito de distribuciones Chi Cuadrado, una correspondiente a cada entero positivo φ. Una distribución Chi Cuadrado correspondiente al número φ se designará por $\chi^2_{(\varphi)}$ y se define por la siguiente función de densidad:

$$f(\chi^2) = \frac{\left[\left(\chi^2\right)^{\left(\varphi/2\right)-1}\right] e^{-\left(\chi^2/2\right)}}{\left(2^{\varphi/2}\right)\left[\Gamma\left(\frac{\varphi}{2}\right)\right]}, \qquad 0 \le \chi^2 \le \infty$$

La función de distribución viene dada por

$$F(x) = P(X \leq x) = \int_0^x f(x)\,dx$$

Si se define ahora lo mismo, pero con $\bar{x}$ en vez de μ, tendremos una χ^2 con φ = n-1 grados de libertad:

$$\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{\sigma^2} = \frac{\hat{\sigma}^2 (n-1)}{\sigma^2} \chi^2_{(n-1)}$$

Recordemos que:

$$\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1} = \hat{\sigma}^2, \quad luego \quad \sum_{i=1}^{n} (x_i - \bar{x})^2 = \hat{\sigma}^2 (n-1)$$

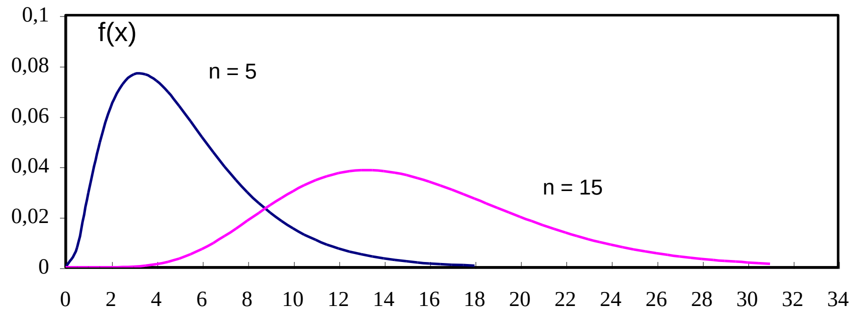
Una variable χ^2 tiene su campo de variación, de 0 (suma de cuadrados) a infinito y su distribución se define completamente por el número de grados de libertad. Si x es $\chi^2_{(\varphi)}$, su media y su varianza, respectivamente son:

$$E\,(\chi^2_{(\varphi)}) = \mu = \varphi \qquad\qquad V\,(\chi^2_{(\varphi)}) = \sigma^2 = 2\varphi$$

Son positivamente asimétricas, pero a medida que aumenta φ, se aproxima a la distribución normal con media φ y desviación $\sqrt{2\varphi}$. En la práctica cuando φ es mayor a 30, se aproxima a la distribución normal.

Gráfica de la función de densidad de la variable χ^2 con 5 y 15 grados de libertad:

Gráfica de la función de densidad de la variable χ^2 con 5 y 15 grados de libertad:



Aplicaciones

Es de utilidad en situaciones en las que se requieren inferencias respecto a la dispersión de una distribución, siendo el caso de una técnica para inferir respecto de la *Varianza* de una distribución en la *Estimación* y en la *Dócima.*

Tablas Usuales

Grados de Libertad	0,995	0,99	0,975	...	0,05	0,025	0,01	0,005
1	7,9	6,6	5,0	...	0,0	0,0	0,0	0,0
2	10,6	9,2	7,4	...	0,1	0,1	0,0	0,0
3	12,8	11,3	9,3	...	0,4	0,2	0,1	0,1
....	....	....	....	...	....	....	....	....
30	53,7	50,9	47,0	...	18,5	16,8	15,0	13,8

En el cuerpo de la tabla se muestran los diferentes valores de la variable para cada combinación de probabilidad y grados de libertad.

Esta tabla acumula desde 0 hasta un valor de la variable.

A los fines de comprender su manejo, plantearemos los siguientes casos:

1- $\Pr(\chi^2_{30} \leq 18{,}5) = \underline{0{,}05}$

A esta probabilidad la encontramos directamente en tabla, pues, como ya dijimos, la misma acumula desde 0 hasta un valor de la variable, en este caso hasta 18,5.

Entonces, ubicamos en primer término los grados de libertad, en este caso 30 y a esta altura buscamos en el cuerpo de la tabla el valor 18,5, la probabilidad que encabeza la columna, es la solicitada.

2- $\Pr(\chi^2_{30} \geq 34{,}8) = 1 - \Pr(\chi^2_{30} \leq 34{,}8) = 1 - 0{,}75 = \underline{0{,}25}$

3- χ^2_{15} *y la probabilidad a la derecha de* χ^2 *es* 0,05.

Nuestros datos pueden expresarse simbólicamente como: $\Pr(\chi^2_{15} \geq \chi^2) = 0{,}05$

Para ingresar en tabla necesitamos la probabilidad desde 0 hasta el valor desconocido, que encontramos de la siguiente manera:

$$\Pr(\chi^2_{15} \leq \chi^2) = 1 - \Pr(\chi^2_{15} \geq \chi^2) = 1 - 0{,}05 = \underline{0{,}95}$$

Con esta información entramos en tabla para encontrar el valor de la variable que surgirá de la intersección entre 15 grados de libertad y la probabilidad de 0,95. Luego $\chi^2 = 25$.

4- χ^2_{15} *y la probabilidad a la izquierda de* χ^2 *es* 0,05.

Nuestros datos pueden expresarse simbólicamente como: $\Pr(\chi^2_{15} \leq \chi^2) = 0{,}05$

Aquí entramos directamente en tabla, buscando el valor de la variable en la intersección entre 15 grados de libertad y 0,05 de probabilidad, encontrando $\chi^2 = 7{,}3$

5- $Pr(\chi^2 \geq 29{,}3) = 0{,}25$

Nuestra incógnita es ahora los grados de libertad, para lo cual necesitaremos:

$$\Pr(\chi^2_\varphi \leq 29{,}3) = 1 - \Pr(\chi^2_\varphi \geq 29{,}3) = 1 - 0{,}25 = 0{,}75$$

Estos valores corresponden a 25 grados de libertad.

1.5.2. *Distribución "t" de Student*

Consideremos una muestra $x_1, x_2, \ldots. x_n$ proveniente de una población con distribución normal cuya media es μ y cuya varianza es σ^2. La variable t es una razón de la variable normal estándar a la raíz cuadrada de un variable Chi Cuadrado dividido por su número de grados de libertad. Es decir que $t_\varphi = \frac{z}{\sqrt{\frac{\chi^2}{\varphi}}}$, donde numerador y denominador son independientes. Hay una distribución t correspondiente a cada entero positivo. La función de densidad para t_φ es como sigue:

$$f(t) = \left(\frac{1}{\sqrt{\varphi\pi}}\right)\left[\frac{\Gamma\left(\frac{\varphi+1}{2}\right)}{\Gamma\left(\frac{\varphi}{2}\right)}\right]\left(1 + \frac{t^2}{\varphi}\right)^{-(\varphi+1)/2}, \quad -\infty \leq t \leq \infty$$

Y la Función De Distribución:

$$P(T \leq t) = \int_{-\infty}^{t} f(t)\,dt$$

Definamos ahora la distribución t del siguiente modo:

1) $z = \frac{x - \mu}{\sigma} \approx N(0,1)$

2) $U = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{\sigma^2} \approx \chi^2_{(n-1)}$ $\quad$ *donde* $\varphi = n - 1$

Entonces, una variable t se define como el cociente entre $Z \sim N(0,1)$ y la raíz cuadrada de

$$U \sim \chi^2_{n-1}$$

Simbólicamente:
$$t = \frac{Z}{\sqrt{\frac{U}{n-1}}} \approx t_{n-1}$$

Reemplazando a Z y U por sus iguales y resolviendo, tenemos:

$$t = \frac{\frac{x-\mu}{\sigma}}{\sqrt{\frac{\sum(\frac{x_i-\bar{x}}{\sigma})^2}{n-1}}} = \frac{\frac{x-\mu}{\sigma}}{\sqrt{\frac{1}{n-1}\sum(\frac{x_i-\bar{x}}{\sigma})^2}} = \frac{\frac{x-\mu}{\sigma}}{\frac{1}{\sigma}\sqrt{\frac{\sum(x_i-\bar{x})^2}{n-1}}}$$

Recordemos que:
$$\sum_{i=1}^{n} \frac{(x_i-\bar{x})^2}{n-1} = \hat{\sigma}^2$$

Reemplazando en el denominador, nos queda que:

$$t = \frac{\frac{x-\mu}{\sigma}}{\frac{1}{\sigma}\sqrt{\hat{\sigma}^2}} = \frac{x-\mu}{\hat{\sigma}} \approx t_{n-1}$$

Aplicación a la media muestral

Sean $x_1, x_2, \ldots. x_n$ una muestra proveniente de una población con distribución normal cuya media es μ y cuya varianza es σ^2.

Formulemos el estadístico "z" de la siguiente manera:

$$z = \frac{\bar{x} \quad \mu}{\sigma_{\bar{x}}} = \frac{\bar{x} \quad \mu}{\frac{\sigma}{\sqrt{n}}} \approx N(0,1),$$

Donde reemplazamos a $\sigma_{\bar{x}}$, por su igual $\frac{\sigma}{\sqrt{n}}$, según demostraremos en la siguiente unidad.

y tomemos:

$$U = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{\sigma^2} \approx \chi^2_{(n-1)}$$

Entonces:

$$t = \frac{\dfrac{\dfrac{\bar{x}-\mu}{\sigma}}{\sqrt{n}}}{\sqrt{\dfrac{1}{n-1}\sum\left(\dfrac{x_i-\bar{x}}{\sigma}\right)^2}} = \frac{\dfrac{(\bar{x}-\mu)\sqrt{n}}{\sigma}}{\dfrac{1}{\sigma}\sqrt{\dfrac{\sum(x_i-\bar{x})^2}{n-1}}} = \frac{(\bar{x}-\mu)\sqrt{n}}{\sqrt{\hat{\sigma}^2}} = \frac{(\bar{x}-\mu)\sqrt{n}}{\hat{\sigma}}$$

Dividiendo numerador y denominador por $\sqrt{n}$, nos queda

$$t = \frac{(\bar{x}-\mu)}{\dfrac{\hat{\sigma}}{\sqrt{n}}} \approx t_{n-1}$$

Observemos que la estructura es la misma que la de la variable z, la diferencia radica en el denominador que en z utiliza la desviación poblacional y en t la desviación muestral corregida.

Es decir, que cuando la muestra es pequeña, el estimador de la varianza, $\hat{\sigma}^2 = \hat{s}^2$, tiene mayor dispersión alrededor del parámetro σ^2, que cuando la muestra es grande, y la variaciones que puede sufrir el estadístico $\dfrac{(\bar{x}-\mu)}{\frac{\hat{\sigma}}{\sqrt{n}}}$ por la variabilidad del denominador, impiden utilizar la distribución normal.

Entonces,

$$z = \frac{\bar{x}-\mu}{\dfrac{\sigma}{\sqrt{n}}} \approx N(0,1)$$, pues supone conocida σ

$$t = \frac{(\bar{x}-\mu)}{\dfrac{\hat{\sigma}}{\sqrt{n}}} \approx t_{n-1}$$, pues se desconoce σ

Cuando el tamaño de la muestra, n, tiende a infinito, es estimador $\hat{\sigma}$ tiende a σ, por lo tanto cuando el tamaño de la muestra llega a ser grande, el estadístico t se aproxima a z.

Propiedades

- $-\infty < t < \infty$, es decir que la variable asume valores entre $-\infty$ hasta ∞.
- Es Unimodal y simétrica respecto a 0.

- Es más aplanada que la distribución normal, pues su varianza es ligeramente superior a 1.
- $E(t) = 0 \qquad y \qquad V(t) = \dfrac{\varphi}{\varphi - 2} \qquad \text{válida para } \varphi \rangle 2$

Aplicaciones

Inferencias sobre la media poblacional, cuando se desconoce la varianza poblacional y el tamaño de muestra es pequeño, tanto en la estimación estadística como en la Docimasia de Hipótesis.

Tablas Usuales

Grados de Libertad	0,75	....	0,9995
1	1,000	....	636,619
....	....	....	....
30	0,683	....	3,646
40			
60			
∞			

Acumula desde -∞ hasta un valor positivo de la variable y los problemas se resuelven de manera idéntica a los vistos en la Distribución Normal, sólo que para entrar en tabla trabajamos con los grados de libertad.

Ejemplos

1-
$\Pr(t_{15} \geq 2{,}602) = 1 - \Pr(t_{15} \leq 2{,}602) = 1 - 0{,}99 = \underline{0{,}01}$

2-

$Pr(\left| t_{18} \right| \geq 2{,}101) = Pr(t_{18} \leq 2{,}101) + Pr(t_{18} \geq 2{,}101) = 1 - Pr(t_{18} \leq 2{,}101) + 1 - Pr(t_{18} \leq 2{,}101) =$

$= 2 \; [1 - Pr(t_{18} \leq 2{,}101)] = 2\,(1 - 0{,}975) = 2\,(0{,}025) = \underline{0{,}05}$

3-
$\Pr(-0{,}706 \leq t_8 \leq 1{,}86) = \Pr(t_8 \leq 1{,}86) - \Pr(t_8 \leq -0{,}706) = \Pr(t_8 \leq 1{,}86) - [1 - \Pr(t_8 \leq 0{,}706)] =$
$= 0{,}95 - (1 - 0{,}75) = \underline{0{,}70}$

4-
$\Pr(\left| t_9 \right| \geq t) = 0{,}05 = \Pr(t_9 \leq -t) + \Pr(t_9 \geq t)$ Y
$\Pr(t_9 \leq t) = 0{,}975 \;\Rightarrow\; t = \left| t_{9/0{,}975} \right| = \underline{\pm\, 2{,}262}$

Modelos Especiales de Probabilidad para Variables Aleatorias Continuas

Variable	*Distribución*	*Función De Densidad*	*Función De Acumulación*	*E( )*	*V( )*	*Ds( )*
(X)	Uniforme Continua	$f_{(x)} = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \forall \text{ otro } x \end{cases}$	$F_{(x)} = \begin{cases} \frac{x-a}{b-a} & a < x < b \\ 0 & x \ge a \\ 1 & x \ge b \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\sqrt{\frac{(b-a)^2}{12}}$
(*T*)	Exponencial	$t\ (x) = \lambda\ e^{-\lambda x}$ para $x \ge 0$	$T(x) = P\ (x \le x_i) = 1-e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\sqrt{\frac{1}{\lambda^2}}$
$x \sim N(\mu, \sigma)$	Normal	$f_{(x)} = \frac{1}{\sigma\sqrt{2\pi}}.e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $\quad -\infty \le x \le \infty$	$F_{(x)} = \int_{-\infty}^{x} f(x)dx$	$\int_{-\infty}^{\infty} x f(x)dx$	$\int_{-\infty}^{\infty} [x - E(x)]^2 f(x)dx$	$\sqrt{\int_{-\infty}^{\infty} [x - E(x)]^2 f(x)dx}$
$z \sim N(0,1)$	Normal	$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$	$F_{(z)} = \int_{-\infty}^{z} f(z)dz$	0	1	1
$\chi^2_{(\varphi)}$	Chi Cuadrado	$f(\chi^2) = \frac{\left[(\chi^2)^{\left(\varphi/2\right)-1}\right] e^{-\left(\chi^2/2\right)}}{\left(2^{\varphi/2}\right)\left[\Gamma\left(\frac{\varphi}{2}\right)\right]},$ $0 \le \chi^2 \le \infty$	$F(x) = P(X \le x) = \int_{0}^{x} f(x)dx$	φ	2φ	$\sqrt{2\varphi}$
$t_{(\varphi)}$	t de Student	$f(t) = \left(\frac{1}{\sqrt{\varphi\pi}}\right)\left[\frac{\Gamma\left(\frac{\varphi+1}{2}\right)}{\Gamma\left(\frac{\varphi}{2}\right)}\right]\left(1+\frac{t^2}{\varphi}\right)^{-(\varphi+1)/2},$ $-\infty \le t \le \infty$	$P(T \le t) = \int_{-\infty}^{t} f(t)dt$	0	$\frac{\varphi}{\varphi-2}$ *válida para* $\varphi \rangle 2$	$\sqrt{\frac{\varphi}{\varphi-2}}$