



Universidad Nacional del Litoral
Facultad de Ingeniería y Ciencias Hídricas

ESTADÍSTICA

Ingeniería Informática

TEORÍA

Mg. Ing. Susana Vanlesberg
Profesor Titular

UNIDAD 1

Probabilidad

INTRODUCCIÓN

El cambio en el ejercicio profesional, determina que los planes de estudio de todas las carreras relacionadas a las ciencias de la tierra, informática, economía, ciencias sociales por citar algunas, incluyan a la Estadística como materia troncal con entidad propia y de verdadera necesidad. Se pretende con ello, que un profesional que se apoye en la cuantificación y en el estudio de lo que observa a diario, entienda y conozca los conceptos básicos de la ciencia que le va a permitir, abandonando conductas pragmáticas, profundizar y comprender el fundamento científico de su área de trabajo.

No se trata de hacer expertos en Estadística. El principal objetivo de los docentes de esta materia se centra en generar una actitud crítica ante cualquier lectura científica, adquirir un lenguaje común con estadísticos y otros profesionales del área. Conocer a priori los pasos y los elementos imprescindibles en cualquier investigación empírica que se apoye en el manejo de grandes volúmenes de datos cuyo propósito final sea condensar dicha información para que pueda ser transmitida o extrapolar las conclusiones a las poblaciones de las que fueron tomadas las medidas.

Aunque no sea su deseo adentrarse en el mundo de la investigación, una parte importante en la transmisión de los nuevos hallazgos y conocimientos de otros colegas de su ámbito profesional, es el lenguaje estadístico.

Es por ello que han de estar absolutamente familiarizados con dicha terminología si se pretende tener una actitud crítica y objetiva ante la lectura de cualquier literatura científica.

En lo que se refiere al modo y la forma de transmitir los conocimientos, la experiencia acumulada a través de los años de docencia y en el área de la investigación nos condiciona a que teoría y práctica avancen de manera simultánea.

¿QUÉ ES LA ESTADÍSTICA?

Cuando coloquialmente se habla de Estadística, se suele pensar en una relación de datos numéricos presentada de forma ordenada y sistemática. Esta idea es la consecuencia del concepto popular que existe sobre el término y que cada vez está más extendido debido a la influencia de nuestro entorno, ya que hoy día es casi imposible que cualquier medio de difusión: periódico, radio, televisión, etc. no nos aborde diariamente con cualquier tipo de información estadística sobre accidentes de tránsito, índices de crecimiento de población, turismo, tendencias políticas, etc. Sólo cuando nos adentramos en un mundo más específico empezamos a percibir que la Estadística no sólo es algo más, sino que se convierte en la única herramienta que, hoy por hoy, permite dar luz y obtener resultados, y por tanto beneficios, en cualquier tipo de estudio, cuyos movimientos y relaciones, por su variabilidad intrínseca, no puedan ser abordadas desde la perspectiva de las leyes deterministas. Podríamos, desde un punto de vista más amplio, definir la Estadística como la ciencia que estudia cómo debe emplearse la información y cómo dar una guía de acción en situaciones prácticas que entrañan incertidumbre.

La Estadística se ocupa de los métodos y procedimientos para recoger, clasificar, resumir, hallar regularidades y analizar los datos, siempre y cuando la variabilidad e incertidumbre sea una causa intrínseca de los mismos. Además se ocupa de realizar inferencias a partir de ellos, con la finalidad de ayudar a la toma de decisiones y en su caso formular predicciones.

Podríamos por tanto clasificar la Estadística en descriptiva, cuando los resultados del análisis no pretenden ir más allá del conjunto de datos, e inferencial cuando el objetivo del estudio es derivar las conclusiones obtenidas a un conjunto de datos más amplio.

Estadística descriptiva: Describe, analiza y representa un grupo de datos utilizando métodos numéricos y gráficos que resumen y presentan la información contenida en ellos.

Estadística inferencial: Apoyándose en el cálculo de probabilidades y a partir de datos muestrales, efectúa estimaciones, decisiones, predicciones u otras generalizaciones sobre un conjunto mayor de datos.

CONCEPTOS BÁSICOS

Individuos o elementos: personas u objetos que contienen cierta información que se desea estudiar.

Población: conjunto de individuos o elementos que cumplen ciertas propiedades comunes.

Muestra: subconjunto representativo de una población.

Parámetro: función definida sobre los valores numéricos de características medibles de una población.

Estadístico: función definida sobre los valores numéricos de una muestra.

En relación al tamaño de la población, ésta puede ser:

Finita, como es el caso del número de mensajes que llegan al servidor de la Facultad en un día.

Infinita, si por ejemplo estudiamos el número de mensajes recibidos por el servidor de la Facultad a lo largo de su vida útil.

Caracteres: propiedades, rasgos o cualidades de los elementos de la población. Estos caracteres pueden dividirse en cualitativos y cuantitativos.

Modalidades: diferentes situaciones posibles de un carácter. Las modalidades deben ser a la vez exhaustivas y mutuamente excluyentes—cada elemento posee una y sólo una de las modalidades posibles.

Clases: conjunto de una o más modalidades en el que se verifica que cada modalidad pertenece a una y sólo una de las clases.

Frecuentemente en la práctica o en el ejercicio de las profesiones, se realizan simplificaciones tales como no considerar la fricción, adoptar un fluido ideal, no considerar gustos o ingresos en la determinación de la demanda de un determinado bien, con el fin de obtener modelos matemáticos relativamente sencillos.

Generalmente estos modelos son determinísticos: un número describe a una variable que se considera independiente, y a través de una fórmula (modelo) se determina un valor específico para la variable dependiente. Para el mismo conjunto de valores de la variable independiente, se obtendrá siempre el mismo conjunto de variables dependientes.

Cuando la incertidumbre en los resultados debe ser considerada explícitamente, sea por variación inherente en la naturaleza, falta de conocimiento de todas las causas y efectos en los sistemas físicos, por conocimiento incompleto del fenómeno que se analiza, falta de datos los modelos que se deberían utilizar son los probabilísticos y se los analiza haciendo uso de las reglas de la Teoría de Probabilidad.

A causa de la existencia de incertidumbre es que los eventos futuros no pueden pronosticarse de una manera exacta. Es más correcto decir que se considera la posibilidad de ocurrencia de un evento y luego determinar su probabilidad.

La Teoría de Probabilidad está relacionada con los experimentos y sus resultados, donde el término "experimento" está usado en un sentido general.

El cálculo de probabilidades suministra las reglas para el estudio de los experimentos aleatorios o de azar, constituyendo la base para la Estadística inductiva o inferencial.

Para trabajar con el cálculo de probabilidades es necesario fijar previamente cierta terminología.

Experimentos y sucesos aleatorios

Nosotros queremos estudiar experimentos que no son determinísticos, pero no estamos interesados en todos ellos. Por ejemplo, no podremos estudiar un experimento del que, por no saber, ni siquiera sabemos por anticipado los resultados que puede dar. No realizaremos tareas de adivinación. Por ello definiremos experimento aleatorio como aquel que verifique ciertas condiciones que nos permitan un estudio riguroso del mismo.

Llamamos experimento aleatorio al que satisface los siguientes requisitos:

- Todos sus posibles resultados son conocidos de antemano.
- El resultado particular de cada realización del experimento es imprevisible.
- El experimento se puede repetir indefinidamente en condiciones idénticas.

Al conjunto de resultados posibles lo denominaremos espacio muestral y lo denotaremos normalmente mediante la letra E o S . Los elementos del espacio muestral se denominan sucesos elementales.

Cualquier subconjunto de E será denominado suceso aleatorio, y se denotará normalmente con las letras A, B, \dots

Por experimento aleatorio o estadístico se entiende al conjunto de acciones que llevan a obtener resultados que pueden variar por causas que no se pueden prever, aunque se realicen en las mismas condiciones. Un experimento puede ser un experimento físico, que se puede repetir cualquier número de veces, o bien simplemente acciones llamadas pruebas, que se repiten en las mismas condiciones.

De acuerdo a los valores que el espacio muestral contenga, puede ser clasificado como:

-finito: si contiene una cantidad dada de valores posibles y es posible listarlos. Por ejemplo, la cantidad de resortes defectuosos en un teclado (de 0 a 121), cantidad de virus detectados en un proceso de corrida de un antivirus.

-infinito numerable: si se sabe que contendrá una gran cantidad de elementos pero posibles de enumerar, y pueden ponerse en correspondencia con el conjunto de los enteros positivos. Por ejemplo: cantidad de CD que se venden en un comercio expendedor de

suministros de computación, cantidad de alumnos ingresantes a la carrera Ing. en Informática en los últimos 10 años.

-infinito o continuo: sus elementos se corresponden con los números reales. Por ejemplo: salario de los operarios de una empresa, Kw consumidos en una industria en los últimos 5 años.

En los dos últimos casos es útil definir el espacio muestral a través de una regla, ya que listar los elementos podría resultar en algunos casos tedioso y en otros imposible.

Por lo general no interesan los elementos que constituyen el espacio muestral individualmente, pero sí un grupo de ellos, en ese caso se estaría analizando a un subconjunto del espacio muestral. Un evento o suceso es un subconjunto de puntos muestrales del espacio muestral S correspondiente a un experimento aleatorio. Un evento simple es aquel que contiene un solo punto muestral y un evento compuesto es el que contiene dos o más puntos elementales.

Decimos que ha ocurrido un suceso cuando se ha obtenido alguno de los resultados que lo forman. **El objetivo de la Teoría de la Probabilidad es estudiar con rigor los sucesos**, que como vemos se pueden enunciar desde el lenguaje común, asignarles probabilidades y efectuar cálculos sobre dichas probabilidades. Observamos que los sucesos no son otra cosa que conjuntos y por tanto, serán tratados desde la Teoría de Conjuntos. Recordamos las operaciones básicas y las dotamos de interpretación para el caso de sucesos. A veces es más sencillo encontrar los elementos que forman el subconjunto contrario al buscado. Suceso complementario de un suceso es aquel formado por los elementos del espacio muestral que no están incluidos en el suceso de interés.

Ejemplo: se considera como experimento observar un objetivo en la pantalla de un visor. El resultado observable es la posición de la mancha luminosa sobre la pantalla, ubicada en un círculo de 10 cm de radio en el sistema de coordenadas cartesianas cuyo origen se halla en el centro de la pantalla. Todos los sucesos observables que interesan en el experimento dado están ligados con el registro de la posición que ocupa la mancha luminosa en la pantalla. Aunque es evidente que no es posible observar físicamente sobre la pantalla un punto, no obstante, semejante procedimiento idealizado para describir un caso elemental simplifica la formalización matemática del experimento en cuestión.

El conjunto S es continuo y puede ser escrito de la siguiente forma:

$$S = \{(x, y) / x^2 + y^2 \leq 100\}$$

1.2 - RELACIONES ENTRE EVENTOS

Los eventos pueden estar relacionados de varias formas:

-Todos los puntos muestrales comunes a ambos eventos se denomina intersección.

-El conjunto que no contiene ningún suceso elemental se denomina evento o suceso imposible (\emptyset).

-Si dos eventos A y B no contienen puntos elementales en común se dice que son mutuamente excluyentes; esta relación puede extenderse a más de dos eventos ($A \cap B = \emptyset$).

-La unión de dos eventos A y B es el evento que contiene los puntos muestrales pertenecientes a uno o a otro o a ambos ($A \cup B$).

Respecto al ejemplo anterior se identifican los siguientes eventos:

A : {el objetivo se encuentra en el primer cuadrante}

$$A = \{(x, y) / x^2 + y^2 \leq 100, x > 0, y > 0\}$$

B : {el objetivo se encuentra en el círculo de 5 cm de

radio cuyo centro coincide con el de la pantalla}

$$B = \{(x, y) / x^2 + y^2 \leq 25\}$$

C : {el objetivo se encuentra en el círculo de 2.5 cm de

radio cuyo centro esta desplazado en 5 cm

a lo largo del eje x en el sentido negativo}

$$C = \{(x, y) / (x + 5)^2 + y^2 \leq 6.25\}$$

Los sucesos A y B son no excluyentes, ya que tienen elementos comunes. Lo mismo sucede con los sucesos B y C, mientras que los sucesos A y C son excluyentes.

1.3 - PROBABILIDAD

Para cada punto en el espacio muestral perteneciente a un experimento es posible asignar un número llamado probabilidad. Existen algunas definiciones de probabilidad que responden a distintas escuelas:

Probabilidad subjetiva

Las probabilidades subjetivas están basadas en las creencias de las personas que efectúan la estimación de probabilidad. Se puede definir como la probabilidad asignada a un evento por parte de un individuo, basada en la evidencia que se tenga disponible. Esa evidencia puede presentarse en forma de frecuencia relativa de presentación de eventos pasados o puede tratarse simplemente de una creencia meditada.

Los tomadores de decisiones pueden hacer uso de cualquier evidencia que tengan a mano y mezclarlas con los sentimientos personales sobre la situación.

Como casi todas las decisiones sociales y administrativas de alto nivel se refieren a situaciones específicas y únicas, los responsables de tomar decisiones hacen un uso considerable de la probabilidad subjetiva.

Probabilidad a priori

La más antigua es la que se originó en los juegos de azar, denominada probabilidad a priori, según Laplace y se basa en el sencillo supuesto de resultados igualmente probables de un experimento. Así, si un suceso definido en el espacio muestral, tiene N resultados posibles igualmente probables y tiene n elementos, la probabilidad de ese suceso se define como el cociente de esas cantidades:

$$P(E) = \frac{n}{N} \quad (1.3.1)$$

siendo n el número de casos favorables o que pertenecen al suceso y N el número de casos igualmente posibles o probables. Esto último hace criticable a la definición, ya que utiliza al término "probable" en la definición de probabilidad.

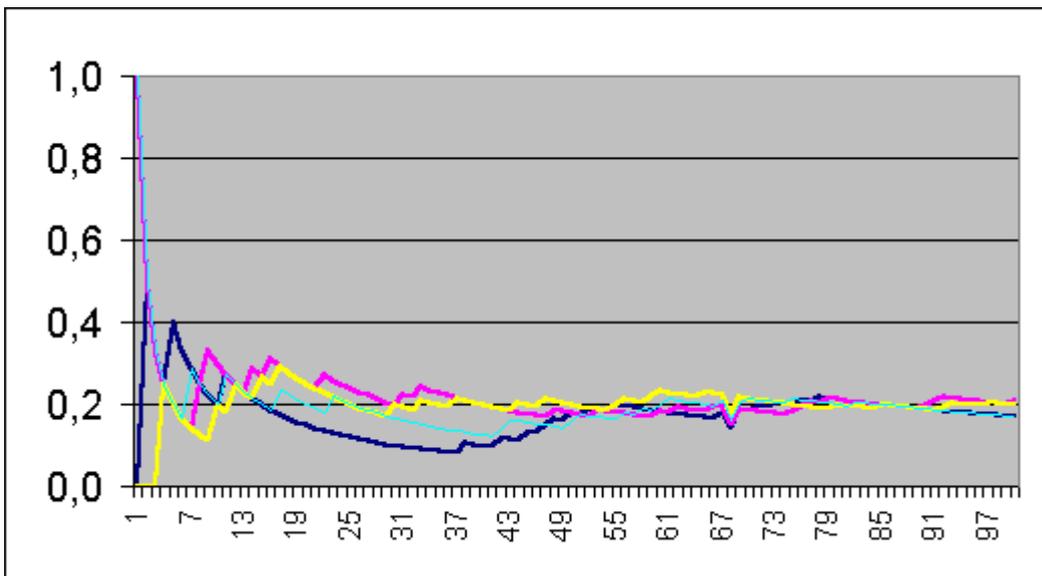
Probabilidad a posteriori o frecuencial

Otra forma de definir a la probabilidad de eventos es a través de la frecuencia relativa, denominada probabilidad a posteriori. Esto es, si un experimento es repetido un número n de veces en condiciones iguales y existen n_1 ($n_1 \leq n$) resultados en los cuales se ha presentado el suceso que interesa, la frecuencia relativa n_1/n puede estimar a su probabilidad. Esta frecuencia relativa se aproxima a un valor constante cuando el número n de pruebas aumenta sin límites, y ese número es la probabilidad del suceso:

$$P(E) = \lim_{n \rightarrow \infty} \frac{n_1}{n} \quad (1.3.2)$$

En la práctica se debe tener un límite y en realidad se toma a la frecuencia relativa como la probabilidad del evento.

En la gráfica siguiente se observa la repetición de una serie de experimentos una determinada cantidad de veces y la estabilización de las frecuencias hacia un valor constante que es su probabilidad.



Probabilidad axiomática

Las dificultades que presenta la definición frecuencial de probabilidad se han resuelto a principios del siglo XX mediante la utilización de una definición axiomática de la probabilidad. La definición, debida al ruso Kolmogorov, es muy parecida a la que damos a continuación.

Sea E el espacio muestral, se define la probabilidad como una aplicación $P: P(E) \rightarrow [0, 1]$ que cumple las siguientes condiciones:

*Axioma I: La probabilidad de un evento es un número mayor o igual que cero y menor o igual que uno.

$$0 \leq P(E) \leq 1$$

*Axioma II: La probabilidad del espacio muestral S es igual a uno.

$$P(S) = 1 \quad (1.3.3)$$

*Axioma III: La probabilidad de un evento el cual es la unión de dos eventos mutuamente excluyentes es la suma de sus probabilidades.

$$A \cap B = \emptyset \quad A \cup B = E$$

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

Esto puede generalizarse como:

$$P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i), \quad \text{con } A_i \cap A_j = \emptyset \quad \text{y } i \neq j \quad (1.3.4)$$

Algunas consecuencias de estos axiomas que tienen gran aplicación práctica:

a – Si \bar{A} es el complemento de A en el espacio muestral S, luego:

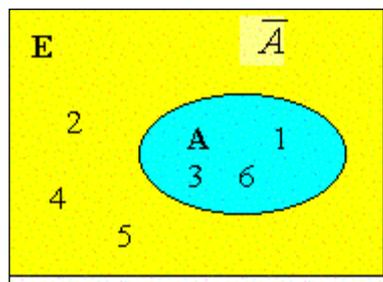
$$\bar{A} \cup A = S$$

$$\bar{A} \cap A = \emptyset$$

$$P(\bar{A} \cup A) = P(S) \quad \text{y } P(S) = 1$$

$$P(\bar{A}) + P(A) = 1$$

$$P(\bar{A}) = 1 - P(A) \quad (1.3.5)$$



b - La probabilidad del suceso imposible es igual a cero. El suceso contrario de S es un suceso imposible, aplicando la consecuencia anterior:

$$P(\bar{S}) = 1 - P(S) = 1 - 1 = 0$$

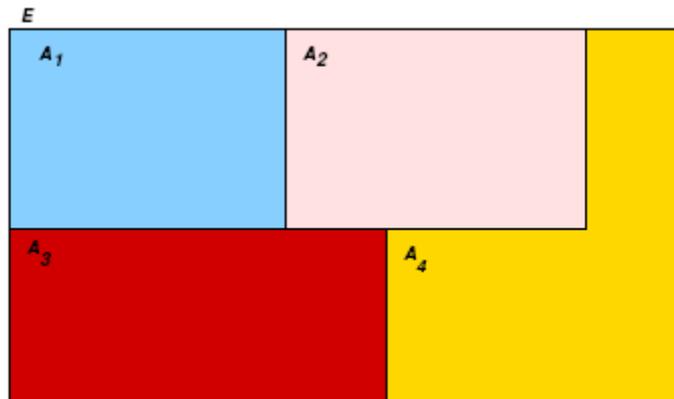
$$P(\emptyset) = 0 \quad (1.3.6)$$

Sistema exhaustivo y excluyente de sucesos

Se dice que la colección $A_1, A_2, \dots, A_n \subset E$ es un sistema exhaustivo y excluyente de sucesos si se verifican las relaciones:

$$\bigcup_{i=1}^n A_i = E$$

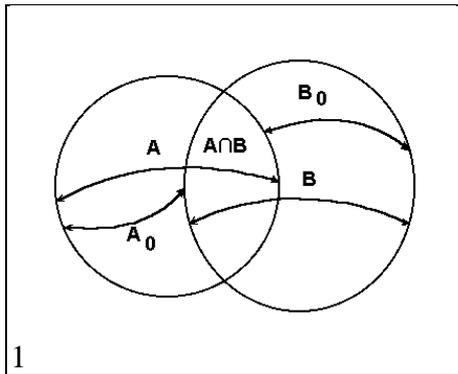
$$A_i \cap A_j = \emptyset \quad \forall i \neq j$$



A_1, A_2, A_3, A_4 forman un sistema exhaustivo y excluyente de sucesos.

1.4 - PROBABILIDAD TOTAL

Teorema. Si dos eventos A y B pertenecen al mismo espacio muestral, la probabilidad de que A o B o ambos ocurran es la suma de sus probabilidades menos la probabilidad de su ocurrencia conjunta.



$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (1.4.1)$$

Esta última expresión se justifica considerando dos sucesos no excluyentes A y B:

$$A = (A \cap B) \cup A_0$$

$$B = (A \cap B) \cup B_0$$

Sus probabilidades, por el Axioma III, son:

$$P(A) = P(A \cap B) + P(A_0)$$

$$P(B) = P(A \cap B) + P(B_0)$$

Luego, la unión de A y B será:

$$A \cup B = A_0 \cup (A \cap B) \cup B_0$$

Entonces, por el Axioma III:

$$P(A \cup B) = P(A_0) + P(A \cap B) + P(B_0)$$

De las expresiones anteriores, es posible obtener $P(A_0)$ y $P(B_0)$ para reemplazarlas en esta última:

$$P(A_0) = P(A) - P(A \cap B)$$

$$P(B_0) = P(B) - P(A \cap B)$$

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

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

con lo que se demuestra la expresión correspondiente a la (1.4.1).

1.5 - PROBABILIDAD COMPUESTA Y CONDICIONAL

Cuando para la ocurrencia de un suceso es necesario que se presenten otros dos sucesos en forma conjunta se habla de probabilidad compuesta. Se deberá analizar si la presentación de uno de estos sucesos previamente condiciona la presentación del otro. Por lo tanto se deberá analizar si se condicionan o no. Un concepto de mucha importancia práctica es el siguiente:

La probabilidad condicional de un evento A dado que el B ha ocurrido se define como el cociente entre la probabilidad de la presentación conjunta de ambos eventos y la probabilidad del suceso B.

$$P(A / B) = \frac{P(A \cap B)}{P(B)} \quad (1.5.1)$$

Esto se refiere a que algunas veces, cuando se calcula la probabilidad de un evento, ya se tiene información; esto reduce el espacio de muestra original a uno de sus subconjuntos, o sea ubicándose en una porción del espacio de muestra y no en cualquier lugar.

Cuando la probabilidad de B es igual a cero, la probabilidad condicional no está definida. De la misma forma:

$$P(B / A) = \frac{P(A \cap B)}{P(A)}; \quad P(A) > 0 \quad (1.5.2)$$

1.5.1 - INDEPENDENCIA

Para poder obtener una expresión general para la probabilidad conjunta o compuesta de dos sucesos, hay que hacer la distinción entre sucesos dependientes e independientes. Se dice que dos eventos son dependientes si la ocurrencia de uno es afectada por la ocurrencia del otro; en caso contrario, los eventos se dicen independientes.

La dependencia e independencia estadística se relacionan con la forma en que se realiza el muestreo.

En general, los sucesos dependientes se relacionan con el muestreo sin reposición, y los independientes con el muestreo con reposición.

El muestreo aleatorio sin reposición es el proceso de selección al azar a partir de una población, de los elementos que conforman la muestra, sin devolverlos a la población antes de volver a extraer.

El muestreo aleatorio con reposición es el proceso de selección con devolución de los elementos antes de realizar una nueva extracción.

La ley que rige la probabilidad de ocurrencia conjunta de sucesos se llama regla de multiplicación de probabilidades y se expresa de la siguiente manera:

La probabilidad de ocurrencia conjunta de dos sucesos A y B está dada por:

$$\begin{aligned} P(A \cap B) &= P(B) \cdot P(A/B) \\ &\quad \text{si son sucesos dependientes} \\ P(A \cap B) &= P(A) \cdot P(B/A) \end{aligned}$$

(1.5.1.1)

$$P(A \cap B) = P(A) \cdot P(B) \quad \text{si son sucesos independientes}$$

(1.5.1.2)

Ejemplos de probabilidad total y compuesta

Las computadoras personales de una red bancaria son de dos tipos: 40% del tipo A y el resto de tipo B. El 10% de las del tipo A sufrieron algún desperfecto durante el último año. Sabiendo que del total de las computadoras el 7% tuvieron desperfecto alguna vez durante el año pasado, determinar la probabilidad de desperfecto en las computadoras de tipo B.

$A : \{\text{Computadora de tipo A}\}$

$$P(A) = 0,40$$

$B : \{\text{Computadora de tipo B}\}$

$$P(B) = 0,60$$

$D : \{\text{desperfecto en cualquier computadora}\}$

$$P(D) = 0,07$$

$D : \{\text{desperfecto en computadora del tipo A}\}$

$$P(D/A) = 0,10$$

$$P(D/B) = ?$$

$$\begin{aligned} P(D) &= P(D \cap A) + P(D \cap B) = \\ &= P(D/A) P(A) + P(D/B) P(B) = \\ 0,07 &= 0,10 * 0,4 + P(D/B) * 0,60 \end{aligned}$$

$$P(D/B) = 0,05$$

Se estima que entre la población del país, el 55% padece de obesidad, el 20% es hipertensa, y el 60% es obesa o hipertensa. ¿Es, de hecho, independiente el que una persona sea obesa de que padezca hipertensión?

Árbol de Probabilidad

Es una herramienta que se utiliza para determinar todos los posibles resultados de experimentos aleatorios.

En el cálculo de la probabilidad se requiere conocer el número de elementos que forman parte del espacio muestral, éstos se pueden determinar con la construcción del diagrama de árbol.

Reglas para calcular probabilidad utilizando diagramas de árbol:

- 1.- la probabilidad de un evento es la multiplicación de las probabilidades de la rama.
- 2.- la probabilidad de varios eventos es la suma de las probabilidades de las ramas con casos favorables al evento.

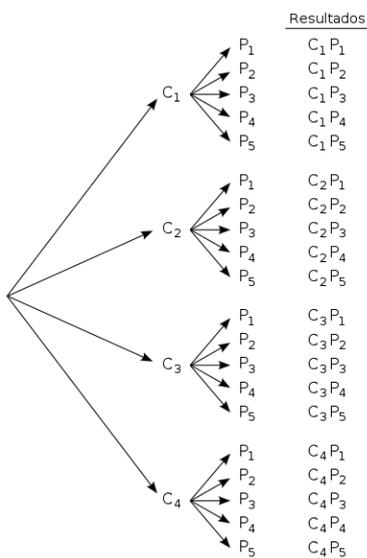
3.- los eventos son mutuamente excluyentes en 2 o más ramas

Para la construcción de un diagrama en árbol se partirá poniendo una rama para cada una de las posibilidades, acompañada de su probabilidad.

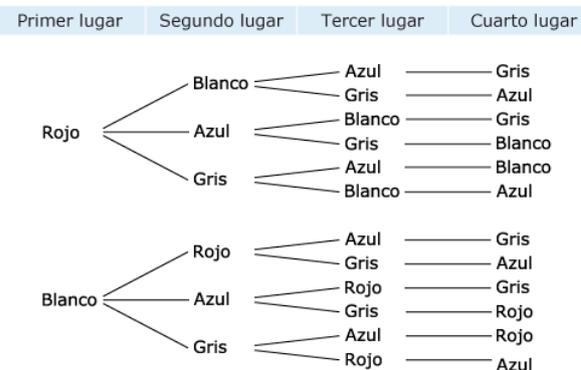
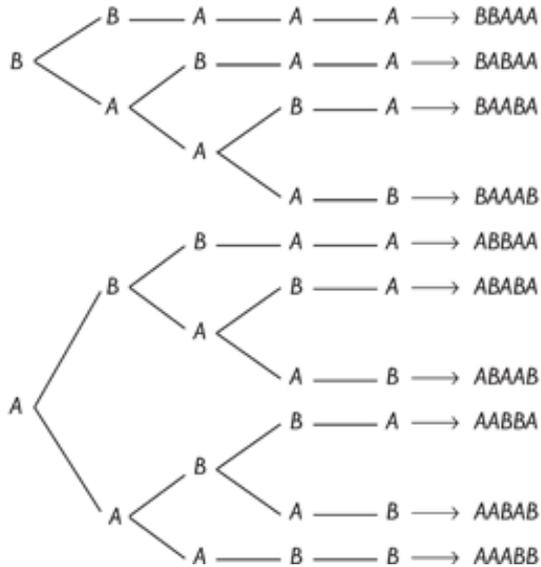
En el final de cada rama parcial se constituye a su vez un nudo del cual parten nuevas ramas, según las posibilidades del siguiente paso, salvo si el nudo representa un posible final del experimento (nudo final).

Hay que tener en cuenta que la suma de probabilidades de las ramas de cada nudo debe dar 1.

Ejemplos:

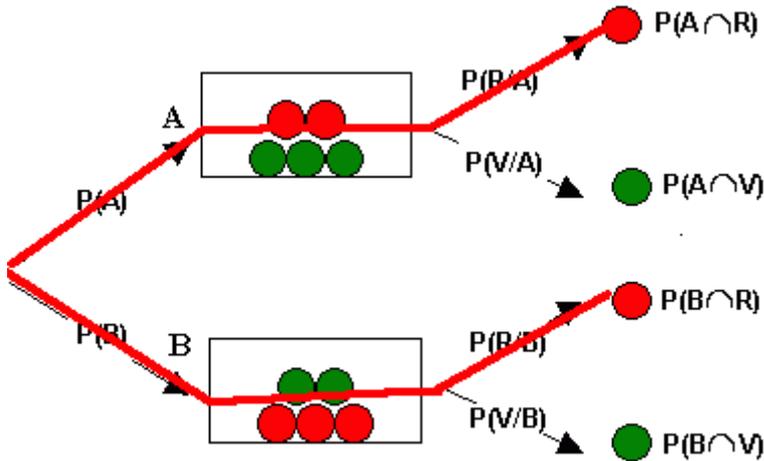


Resultados	
C ₁	P ₁
C ₁	P ₂
C ₁	P ₃
C ₁	P ₄
C ₁	P ₅
C ₂	P ₁
C ₂	P ₂
C ₂	P ₃
C ₂	P ₄
C ₂	P ₅
C ₃	P ₁
C ₃	P ₂
C ₃	P ₃
C ₃	P ₄
C ₃	P ₅
C ₄	P ₁
C ₄	P ₂
C ₄	P ₃
C ₄	P ₄
C ₄	P ₅

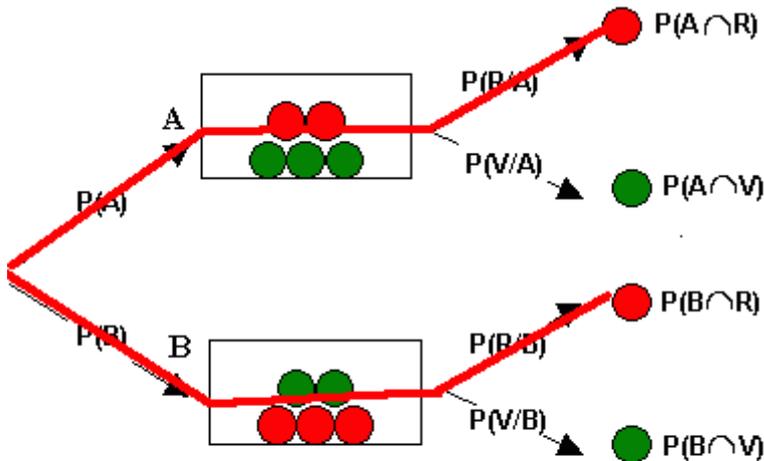


1.6 - TEOREMA DE BAYES (Probabilidades a posteriori)

Considerar el experimento de dos cajas A y B, que contienen tubos verdes y rojos:



Si se sabe que ha salido un tubo rojo, los caminos posibles en el árbol de probabilidades, quedan reducidos a dos, los señalados en rojo en el gráfico anterior; se tiene que reasignar probabilidades, todos los caminos que terminan en tubo verde, deberán tener probabilidad 0. ¿Cómo se asignan probabilidades a los caminos que conducen a tubo rojo?



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

$$P(A/R) = \frac{P(A \cap R)}{P(A \cap R) + P(B \cap R)} = \frac{P(A) \cdot P(R/A)}{P(A) \cdot P(R/A) + P(B) \cdot P(R/B)}$$

Este teorema tiene gran importancia a la hora de tomar decisiones ya que permite incorporar nueva información para mejorar las estimaciones de probabilidad ya realizadas o conocidas, esas probabilidades se denominan probabilidades revisadas o a posteriori.

Resumiendo supóngase un conjunto de eventos mutuamente excluyentes (G_1, G_2, \dots, G_n) con probabilidades mayores que cero, respecto a un experimento aleatorio y que pueden considerarse como un conjunto de hipótesis acerca del suceso que interesa (E) y pertenecientes al espacio muestral. Si el experimento se ha realizado y se conoce que el suceso E ha ocurrido, la probabilidad a posteriori, de que se realice la hipótesis G_i , a condición de que E haya ocurrido, se obtiene a través de la fórmula de Bayes.

$$P(G_i/E) = \frac{P(G_i \cap E)}{P(G_1 \cap E) + P(G_2 \cap E) + \dots + P(G_n \cap E)}$$

$$P(G_1 \cap E) + P(G_2 \cap E) + \dots + P(G_n \cap E) = P(E)$$

$$P(E/G_i) = \frac{P(E \cap G_i)}{P(G_i)} \quad P(E \cap G_i) = P(G_i)P(E/G_i)$$

$$P(G_i/E) = \frac{P(G_i) \cdot P(E/G_i)}{\sum_{i=1}^n P(G_i) \cdot P(E/G_i)}, \text{ con } i = 1, 2, \dots, n \quad (1.6.1)$$

De esta manera la ocurrencia de E condiciona la información previa $P(G_i)$. Estas probabilidades $P(G_i)$ son llamadas probabilidades a priori, ya que son las probabilidades de las causas G_i antes de observar E . De la misma manera $P(G_i/E)$ es llamada probabilidad a posteriori ya que se refiere al estado de conocimiento de G_i después de la ocurrencia de E . Se ve que la fórmula de Bayes permite con información experimental nueva, actualizar la que existe relacionada a un evento.

Ejemplo

En una compañía compran componentes electrónicos de dos proveedores. 60% son comprados en May Electric, y el resto en Harmon Products. El nivel de calidad de May Electric es mejor que el de Harmon Products. 5% de los aparatos comprados en May Electric necesitan mantenimiento adicional, mientras que 8% de los de Harmon Products lo necesitan.

Se seleccionó un componente al azar y se lo encontró defectuoso. ¿Determine la probabilidad de que haya sido comprado en Harmon Products?

$A: \{ \text{El componente se reconoce como defectuoso} \}$

$H_1: \{ \text{El componente se compró a May Electric} \}$

$H_2: \{ \text{El componente se compró a Harmon products} \}$

$$P(H_1) = 0.6 \quad P(H_2) = 0.4$$

$$P(A / H_1) = 0.05 \quad P(A / H_2) = 0.08$$

$$P(H_2 / A) = \frac{P(H_2)P(A / H_2)}{P(A)} =$$

$$P(H_2 / A) = \frac{0.40(0.08)}{0.60(0.05) + 0.40(0.08)} = 0.5161$$

Este fenómeno tiene aplicaciones fundamentales en Ciencia: Cuando se tienen por ejemplo dos teorías científicas diferentes, T1 y T2, que pretenden explicar cierto fenómeno, y a las que asociamos unas probabilidades a priori de ser ciertas, P[T1] , P[T2] podemos llevar a cabo la experimentación que se considere más conveniente, para una vez obtenido el cuerpo de evidencia, B, calcular como se modifican las probabilidades de cada teoría mediante el teorema de Bayes: P[T1|B] , P[T2|B]. Así la experimentación puede hacer que una teoría sea descartada si P[Ti|B] \approx 0 o reforzada si P[Ti|B] \approx 1.



Universidad Nacional del Litoral
Facultad de Ingeniería y Ciencias Hídricas

ESTADÍSTICA

Ingeniería Informática

TEORÍA

Mg.Ing. Susana Vanlesberg
Profesor Titular

UNIDAD 2

VARIABLES ALEATORIAS

Al hablar de sucesos elementales, se definieron casi exclusivamente en forma cualitativa (buenos y malos; defectuosos y no defectuosos; varón y mujer, etc.); pero muchas veces es conveniente cuantificar resultados. Es más, la mayor parte de los problemas que se plantean en la práctica tienen que ver con cantidades medibles. Esta cuantificación se realiza a través de una variable que será de tipo aleatoria.

Formalmente *variable aleatoria* es aquella función definida en el espacio muestral que asocia a cada punto muestral un número perteneciente a los reales. Al realizar el experimento aleatorio la variable toma un solo valor.

Clasificación

Al observar un determinado fenómeno es sencillo deducir que existen distintos tipos de variables aleatorias. Si el fenómeno estudiado u observado es tal que puede ser cuantificado por valores contables o enteros, entonces la variable aleatoria se dice **discreta**. El conjunto de los valores posibles de esta variable puede ser finito o infinito numerable. Como ejemplo, cantidad de procesadores Pentium Intel defectuosos en una partida.

Si el fenómeno puede ser cuantificado a través de cualquier número real, la variable se dirá **continua**. Por ejemplo monto de ganancias anuales de una empresa.

2.1 - Distribución de Probabilidades

Si se considera la variable aleatoria "cantidad de pulsos telefónicos registrados en un nodo Internet en un mes determinado" se verifica que hay cierta cantidad de resultados posibles: pueden haber 200, 300, 1500 etc. También puede asegurarse que hay una cierta probabilidad que sean 300 pulsos, 1500, etc.

Existe una asignación de probabilidades a cada uno de los posibles valores que puede tomar una variable aleatoria. Esa correspondencia entre los valores de la variable aleatoria y su probabilidad de ocurrencia es lo que se denomina **Distribución de Probabilidades**.

Distribución de Probabilidades

El comportamiento de una variable aleatoria es descrito por sus leyes de probabilidad, las cuales pueden ser caracterizadas de distintas formas:

- a- la forma más común es a través de la distribución de probabilidades, el caso más sencillo es a través de una lista de los valores que la variable aleatoria puede tomar y sus respectivas probabilidades (sólo posible en el caso de variable aleatoria discreta);
- b- a través de gráficos
- c- a través de modelos matemáticos que representen la ley.

2.1.1- Variable Discreta. Función Masa de probabilidad o Función de Cuantía.

Esta Función es la forma matemática de expresar la correspondencia entre los valores de la variable y sus probabilidades:

$$f(x) = P(X = x) \quad (2.1)$$

Para satisfacer los axiomas de la Teoría de Probabilidad y ser una función de probabilidad, deberá cumplir con los siguientes requisitos:

- a- $f(x_i)$ debe tener un valor numérico para todos los posibles valores de la variable aleatoria.
- b- $0 \leq f(x_i) \leq 1$ para cualquier valor de x
- b- $\sum f(x_i) = 1$ para todos los valores de x

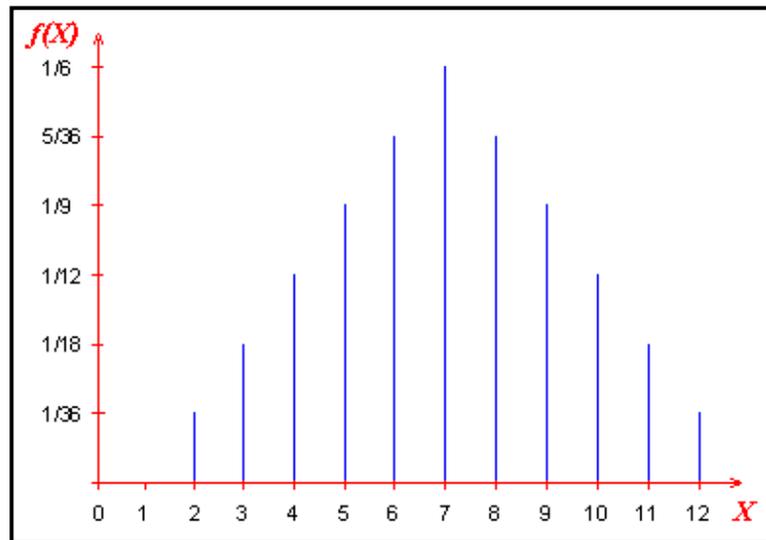
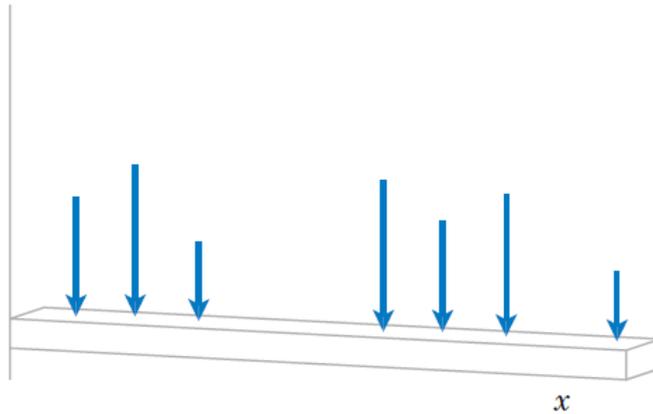


Fig. N° 1- Función de cuantía

Puede hacerse una semejanza con mecánica considerando una masa total unitaria distribuida sobre el eje de las x , con la masa ubicada en cada uno de los puntos y proporcional en altura a la probabilidad de que la variable tome ese valor.



Otra forma de describir la distribución de probabilidades de una variable aleatoria es a través de una función denominada **función acumulativa** o **función de distribución**. Esta función $F(x)$ es simplemente la probabilidad que la variable aleatoria tome valores menores o iguales que un valor determinado:

$$F(x) = P(X \leq x) \tag{2.2}$$

Para el caso de variable aleatoria discreta esta función es la suma de los valores de la función masa de probabilidad evaluada en todos aquellos valores menores o iguales que x , que la variable puede tomar:

$$F(x) = \sum_{\forall x_i \leq x} f(x_i) \tag{2.3}$$

La gráfica de esta función es de tipo escalonada, ya que experimenta saltos en los distintos x_i , iguales a la probabilidad de que la variable aleatoria tome esos valores. Entre dos valores de la variable la probabilidad permanece constante:

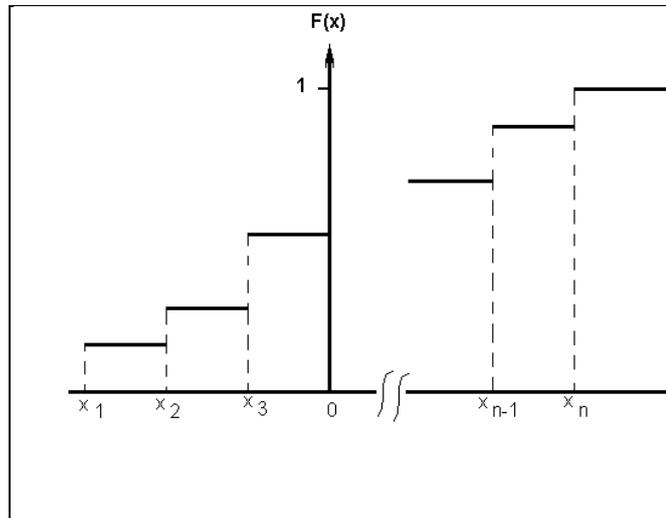
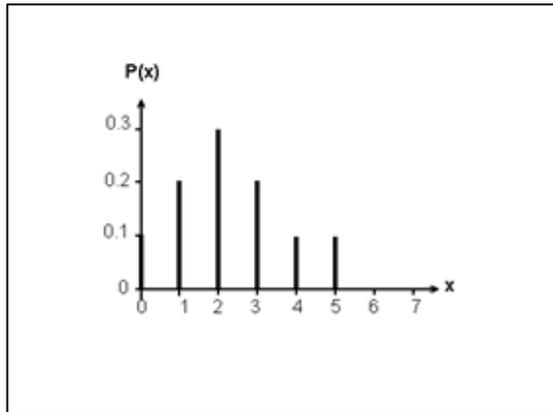


Fig. N° 2- Función de distribución

Ejemplo: considérese la variable aleatoria X N° de bultos de carga que llegan a una fábrica en una semana determinada del mes de mayo. Las probabilidades para los distintos casos son:

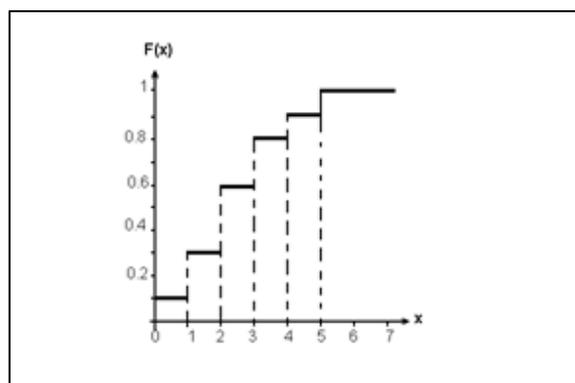
$$p(x) = \begin{cases} 0.1 & x = 0 \\ 0.2 & x = 1 \\ 0.3 & x = 2 \\ 0.2 & x = 3 \\ 0.1 & x = 4 \\ 0.1 & x = 5 \\ 0 & x = 6, 7, \dots \end{cases}$$

Su gráfica correspondiente es la siguiente:



La función de distribución y su gráfica son las siguientes:

$$F(x) = P(X \leq x) = \begin{cases} 0, & x < 0 \\ 0.1, & 0 \leq x < 1 \\ 0.3, & 1 \leq x < 2 \\ 0.6, & 2 \leq x < 3 \\ 0.8, & 3 \leq x < 4 \\ 0.9, & 4 \leq x < 5 \\ 1, & x \geq 5 \end{cases}$$



Se puede decir que la probabilidad de recibir 2 bultos o menos en la fábrica $F(2) =$

$$0.6 = P(x=0)+P(x=1)+P(x=2) = 0.1 + 0.2 + 0.3 = 0.6.$$

Variable aleatoria continua. Función de densidad.

El problema de especificar la distribución de probabilidades y de esta manera la ley para una variable aleatoria continua es sencillo de derivar realizando el siguiente razonamiento:

Si el eje x se divide en un gran número de intervalos infinitesimales, cada uno de longitud dx , es posible definir una función $f(x)$, tal que la probabilidad de que x esté en un intervalo $(x, x+dx)$ es $f(x)dx$. Esta función se denomina *función de densidad de probabilidad* o, simplemente, *función de densidad*.

Ya que ocurrencias en distintos intervalos son eventos mutuamente excluyentes, se deduce que la probabilidad de que una variable aleatoria tome valores en un intervalo de longitud finita, es la suma de probabilidades en los intervalos infinitesimales, o sea, es $\int f(x)dx$ sobre el intervalo de interés.

De esta manera, el área bajo la curva de esta función en un intervalo, representa la probabilidad de que la variable tome valores en el intervalo referido:

$$P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} f(x)dx \quad (2.4)$$

La probabilidad de que una variable aleatoria continua tome un valor específico es 0, ya que la longitud del intervalo dx desaparecería:

$$P(X = x_1) = \int_{x_1}^{x_1} f(x)dx = 0$$

Como consecuencia de esto, las siguientes expresiones son equivalentes:

$$P(a \leq X \leq b) \quad P(a < X < b) \quad P(a \leq X < b) \quad P(a < X \leq b)$$

El valor de $f(x)$ en sí mismo no es una probabilidad, sino solamente la medida de la densidad de probabilidad en un intervalo. Debe cumplir con dos condiciones:

$$f(x) \geq 0$$

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

Puede hacerse una analogía con mecánica, diciendo que se tiene una masa total unitaria distribuida en forma continua en el intervalo en el que toma valores la variable.

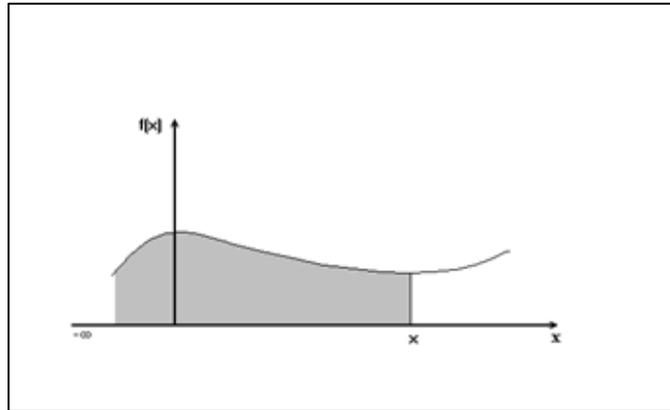
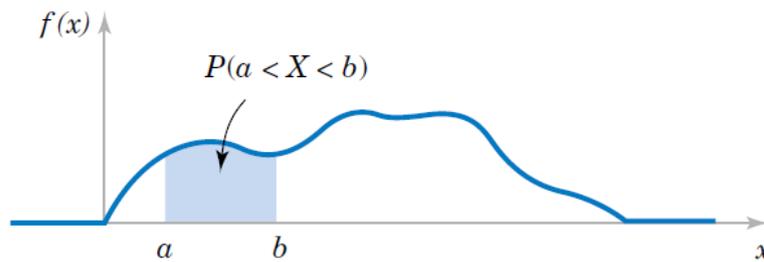


Fig. N° 3 - Función de densidad



Función de Distribución

Se define la función de distribución acumulativa de la v.a. X , $F(x)$, como la probabilidad de que la v.a. continua X tome valores menores o iguales a x , es decir:

$$F(x) = P(X \leq x) \quad (2.5)$$

$$F(x) = \int_{-\infty}^x f(u) du \quad (2.6)$$

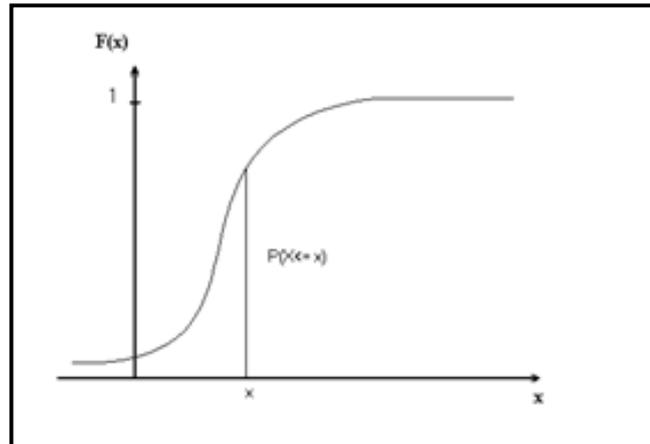


Fig. N° 4 - Función de distribución

Consideramos una masa unidad distribuida sobre el intervalo $(-\infty, +\infty)$, y la función de distribución $F(x)$ para cada valor x de la v.a. da la cantidad de masa que hay en el intervalo $(-\infty, x)$, es decir, la masa que hay en el punto x y a la izquierda de x , aunque dará igual considerar el intervalo $(-\infty, x)$.

Propiedades de la Función de Distribución:

$$F(-\infty) = 0, \quad F(x) = \lim_{x \rightarrow -\infty} \int_{-\infty}^x f(u) du = 0$$

$$F(\infty) = 1, \quad F(x) = \lim_{x \rightarrow \infty} \int_{-\infty}^x f(u) du = 1$$

Es una función monótona creciente:

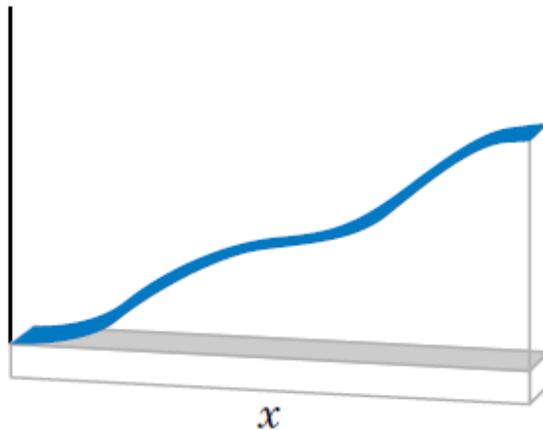
Si $x < x+m$ entonces $F(x) \leq F(x+m)$.

$$F(x+m) = F(x) + P(x < X \leq x+m)$$

Esto permite obtener la probabilidad en un intervalo:

$$F(x+m) - F(x) = P(x < X \leq x+m)$$

$$P(a < X < b) = F(b) - F(a)$$



Puede establecerse la relación entre ambas funciones, considerando un Δx arbitrariamente pequeño y evaluar la función acumulativa en los extremos de ese intervalo: $x + \Delta x$ y x , haciendo tender a cero la amplitud Δx :

$$\lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x}$$

$$\lim_{\Delta x \rightarrow 0} \frac{P(X \leq x + \Delta x) - P(X \leq x)}{\Delta x}$$

$$F'(x) = f(x) = \frac{P(x < X \leq x + \Delta x)}{\Delta x}$$

luego:

$$F(x) = \int_{-\infty}^x f(t) dt \quad (2.7)$$

2.2 - Variables aleatorias bidimensionales

Muchas veces es de interés observar en forma simultánea dos o más variables, por ejemplo cantidad de equipos de computación fabricados por una compañía y cantidad de operarios que trabajan en ella.

Se puede definir variable aleatoria bidimensional como el par de números que expresa el resultado de un experimento combinado y esto puede ser representado en el plano x y ; el recorrido o rango será ahora un subconjunto del plano x y .

También corresponde hacer la distinción entre variable bidimensional discreta y bidimensional continua. El comportamiento conjunto de las variables aleatorias bidimensionales es descrito por su ley de probabilidad conjunta. Se puede también hacer la analogía con mecánica considerando que es la distribución de una unidad de masa sobre el plano x y : en forma continua sobre todo el plano x y si la variable aleatoria bidimensional es continua y concentrada en un número finito o infinito numerable de puntos con masa igual a $p(x_i, y_j)$ en cada punto (x_i, y_j) si la variable bidimensional es discreta.

2.2.1- Variable aleatoria bidimensional discreta. Función masa de probabilidades conjuntas

Esta función es la correspondencia entre cada par de puntos y su probabilidad de ocurrencia. Analíticamente se expresa por:

$$f(x, y) = P(X = x; Y = y) \quad (2.8)$$

ya que es una probabilidad debe cumplir con las siguientes condiciones :

$$f(x, y) \geq 0$$

$$\sum_{\forall x_i} \sum_{\forall y_j} f(x_i, y_j) = 1, \quad \text{con } i = 0, 1, \dots, m \quad j = 0, 1, \dots, n$$

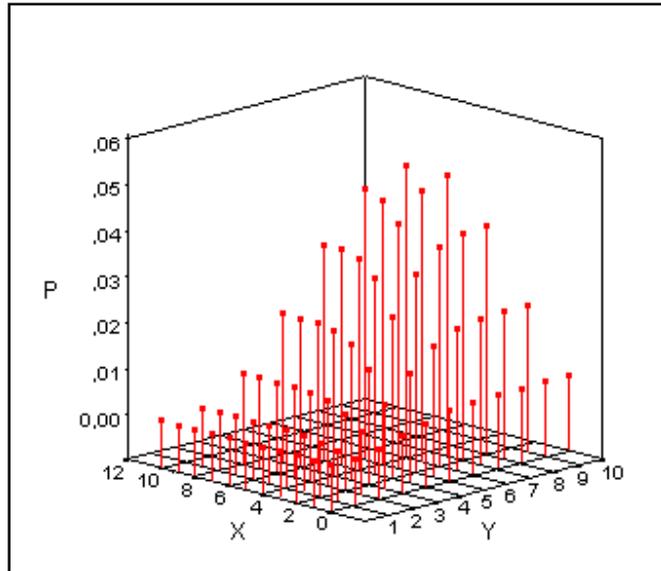


Fig. N° 5- Función de cuantía conjunta

Esta distribución conjunta puede presentarse también a través de una tabla a doble entrada, considerando cada valor de cada variable en los extremos de la tabla y en el cuerpo de ella las probabilidades de presentación conjunta (para cada par de valores). Esto sólo puede hacerse si la variable aleatoria bidimensional es discreta.

	x	x1	x2	...	xm	Distrib. Marginal de Y
y						
y1						
y2			f(x2,y2)			
...						
yn						
Distrib. Marginal de X						1

$$F(x, y) = P(X \leq x; Y \leq y) = \sum_{\forall x_i \leq x} \sum_{\forall y_j \leq y} f(x_i, y_j), i = 0, 1, \dots, m \quad j = 0, 1, \dots, n \quad (2.9)$$

Funciones marginales. Funciones masa de probabilidad marginales

El comportamiento de una variable en particular sin considerar la otra, se describe a través de las funciones marginales. Se las obtiene a partir de la función masa de probabilidades conjuntas, por darle todos los valores de su recorrido a la variable que no interesa considerar. En forma analítica se expresa a través de:

$$f(x) = P(X = x) = \sum_{\forall y_j} f(x, y_j) \quad \text{con } j = 0, 1, \dots, n \quad (2.10)$$

$$f(y) = P(Y = y) = \sum_{\forall x_i} f(x_i, y) \quad \text{con } i = 0, 1, \dots, m$$

Funciones acumulativas marginales

$$F(x) = P(X \leq x) = \sum_{\forall x_i \leq x} f(x_i) = \sum_{x_i \leq x} \sum_{y_j=0}^n f(x_i, y_j) \quad (2.11)$$

$$F(y) = P(Y \leq y) = \sum_{y_j \leq y} f(y_j) = \sum_{x_i=0}^m \sum_{y_j \leq y} f(x_i, y_j)$$

Funciones condicionales

Pueden obtenerse a partir de la distribución conjunta, es aquella que se deriva al evaluar las probabilidades de presentación de una variable conociendo que la otra variable ha ocurrido con un valor particular.

Para un valor particular de Y, $Y=y_0$ la función de cuantía de x condicionada a Y está dada por :

$$f(x/y) = P(X = x | Y = y_0) = \frac{P(X = x \wedge Y = y)}{P(Y = y_0)}$$

$$f(x/y) = \frac{f(x, y)}{\sum_{\forall x_i} f(x_i, y)} = \frac{f(x, y)}{f(y)}, \quad \text{con } f(y) \neq 0 \quad (2.12)$$

Esta función de cuantía condicional, por ser una función de probabilidad debe cumplir con las siguientes condiciones:

$$a - 0 \leq f(x/y) \leq 1$$

$$b - \sum_{\forall x_i} f(x/y) = 1$$

La distribución condicional de Y dado un valor particular de X se define de forma similar:

$$f(y/x) = \frac{f(x, y)}{f(x)}, \text{ con } f(x) \neq 0 \quad (2.13)$$

Variable aleatoria bidimensional continua

Las funciones asociadas con variables aleatorias bidimensionales continuas son análogas a las del caso discreto, pero la función de cuantía será reemplazada por la función de densidad conjunta, siendo $f(x,y) \Delta x \Delta y$ la probabilidad de que la variable (x,y) caiga en el intervalo $x, x+\Delta x$; $y, y+\Delta y$.

La probabilidad de ocurrencia conjunta de x e y en alguna región del espacio muestral se determina por integrar la función de densidad conjunta en esta región:

$$P(x_1 \leq X \leq x_2; y_1 \leq Y \leq y_2) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f(x, y) dx dy \quad (2.14)$$

Esto es igual al volumen debajo de la función de densidad conjunta $f(x,y)$, sobre la región limitada por x_1, x_2 e y_1, y_2 .

La función de densidad conjunta debe cumplir con las siguientes condiciones:

$$f(x, y) \geq 0$$

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

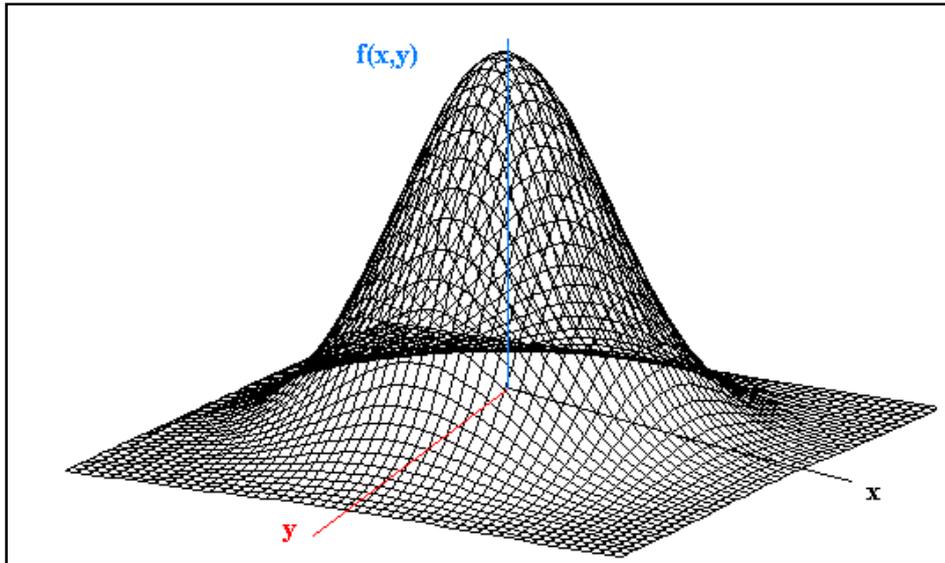


Fig. N° 6- Función de densidad conjunta

La función de distribución conjunta se define como:

$$F(x, y) = P(-\infty \leq X \leq x; -\infty \leq Y \leq y)$$

$$F(x, y) = P(X \leq x; Y \leq y)$$

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(t, l) dt dl$$

(2.15)

Propiedades:

Son similares al caso unidimensional:

$$F(-\infty; -\infty) = 0$$

$$F(\infty; \infty) = 1$$

$$F(-\infty; y) = 0$$

$$F(x; -\infty) = 0$$

ya que $F(X, Y)$ es una probabilidad deberá tomar valores entre 0 y 1.

A partir de esta función, se puede obtener la de densidad conjunta por derivar parcialmente:

$$\frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y) \quad (2.16)$$

Funciones de densidad marginales

Como en el caso discreto, para estudiar el comportamiento de una de las variables en particular se elimina la otra, integrando sobre todos sus valores:

$$f(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad (2.17)$$

$$f(y) = \int_{-\infty}^{\infty} f(x, y) dx \quad (2.18)$$

Funciones de distribución marginales

$$F(x) = P(X \leq x) = \int_{-\infty}^x \int_{-\infty}^{\infty} f(x, y) dy dx = \int_{-\infty}^x f(t) dt$$

$$F_1(x) = F_{x,y}(x, \infty) \quad (2.19)$$

De lo anterior se deduce que:

$$f(x) = \frac{\partial}{\partial x} [F(x, \infty)] = \frac{\partial}{\partial x} [F(x)] \quad (2.20)$$

De forma similar:

$$F_2(y) = P(Y \leq y) = \int_{-\infty}^y \int_{-\infty}^{\infty} f(x, y) dx dy = \int_{-\infty}^y f(s) ds \quad (2.21)$$

$$F(y) = F_{x,y}(\infty, y)$$

Se deduce que:

$$f(y) = \frac{\partial}{\partial y} F(\infty, y) = \frac{\partial}{\partial y} [F(y)] \quad (2.22)$$

Gráficamente:

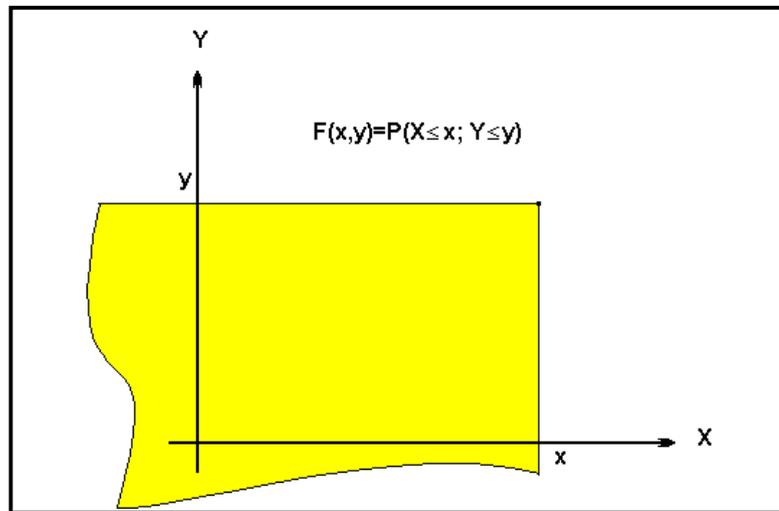


Fig. N° 7- Función de distribución conjunta

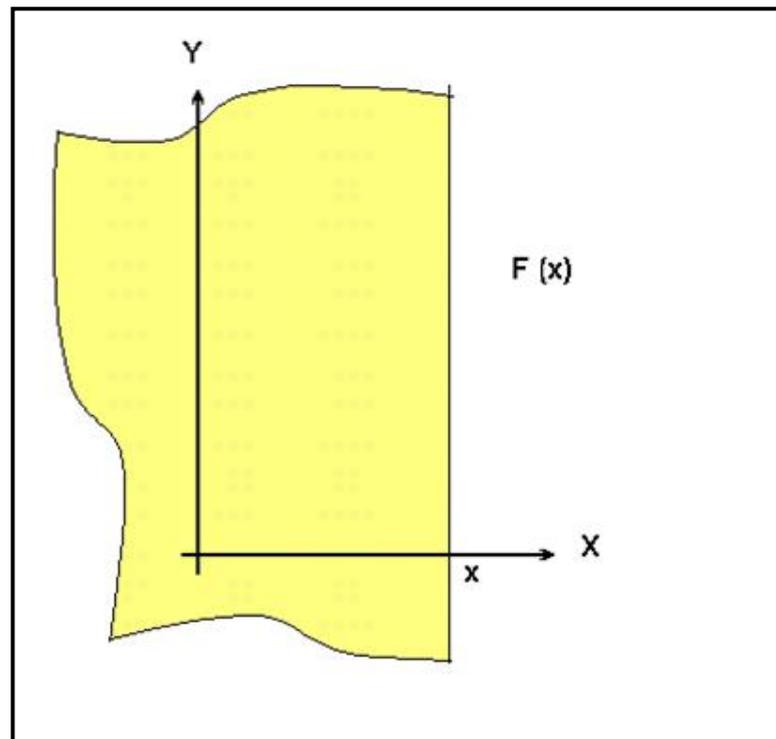


Fig. N° 8 -Función de distribución marginal de X

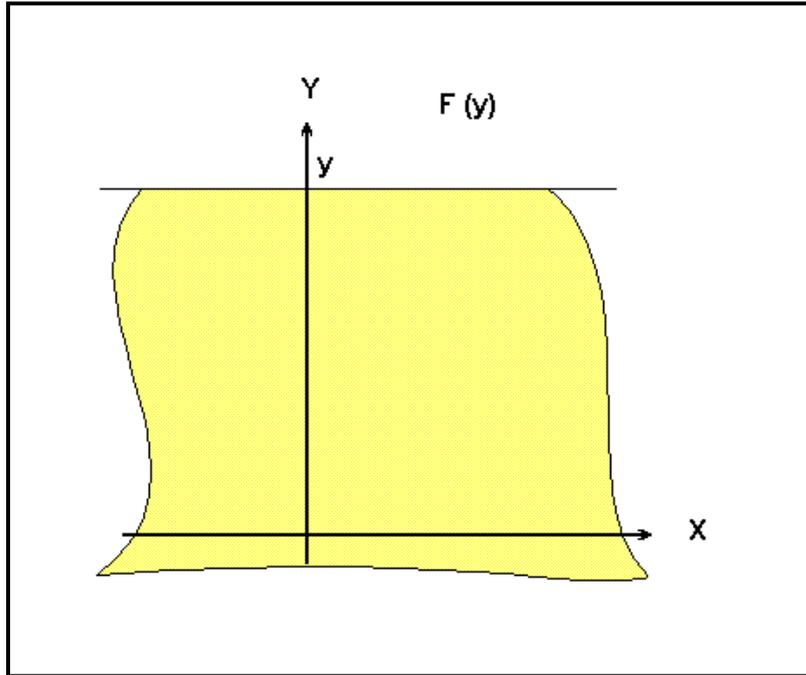


Fig. N° 9 -Función de distribución marginal de Y

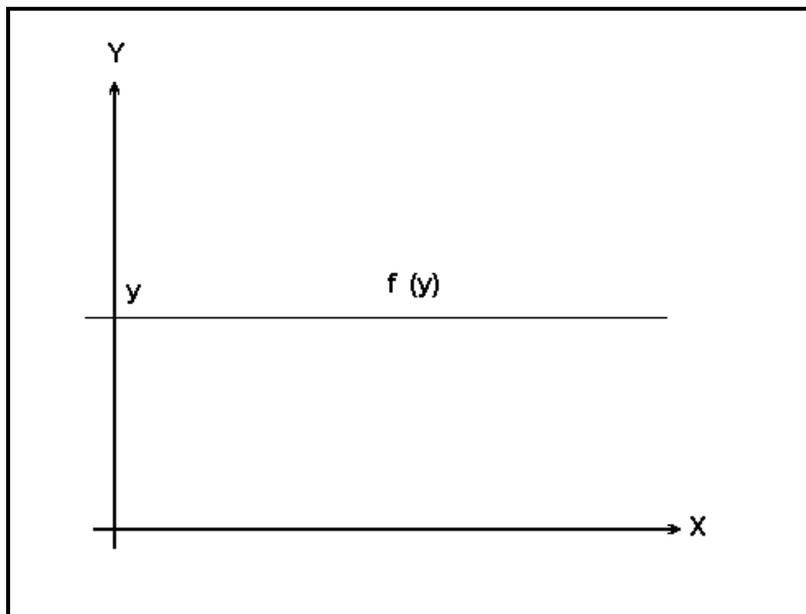


Fig. N° 10 -Función de densidad marginal de Y

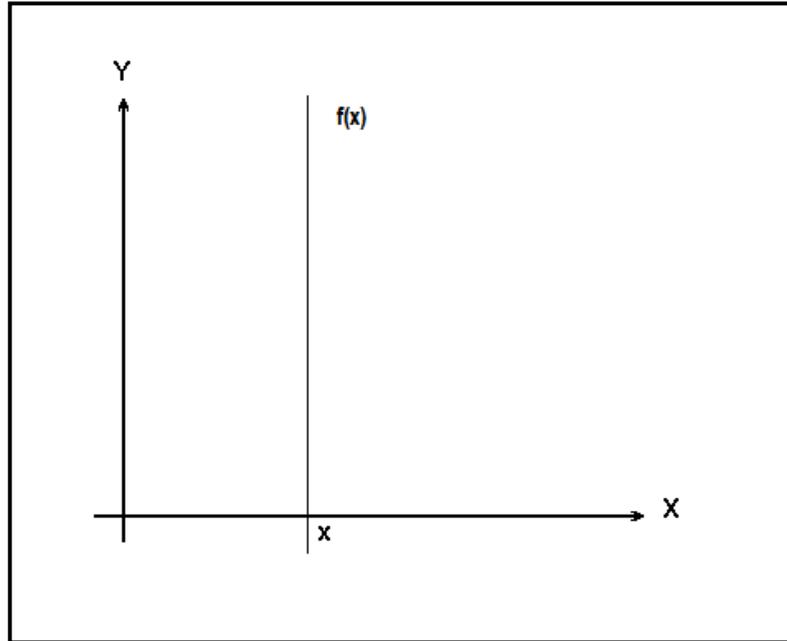


Fig. N° 11 -Función de densidad marginal de X

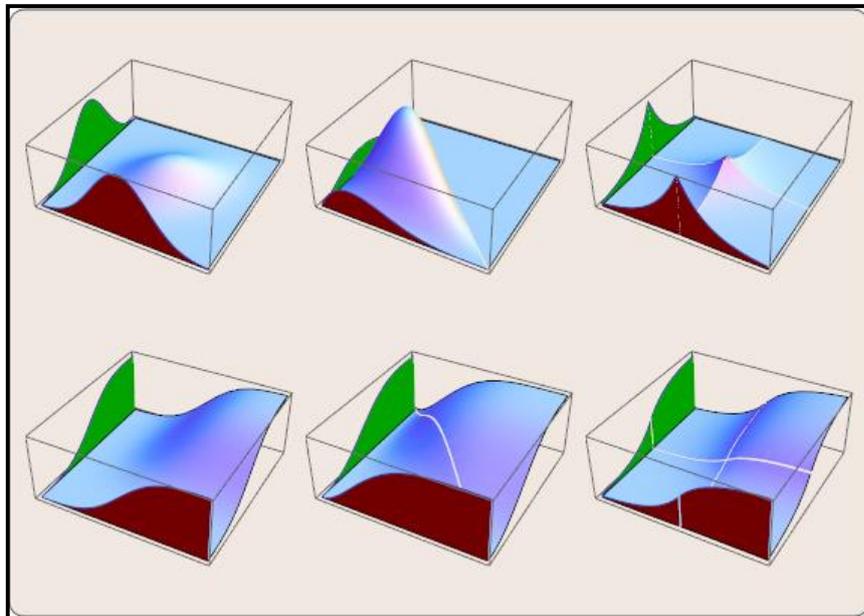


Fig. N° 12 –Funciones marginales

Si interesa obtener la probabilidad en un intervalo para una de las variables puede obtenerse a partir de la función conjunta o a partir de las marginales:

$$P(a < X < b) = \int_a^b f(x) dx$$

$$P(a < X < b) = \int_a^b \int_{-\infty}^{\infty} f(x, y) dy dx$$

(2.23)

Funciones de probabilidad condicionales.

Recordando el concepto de sucesos dependientes:

$$P(A/B) = \frac{P(AB)}{P(B)} \quad P(B/A) = \frac{P(AB)}{P(A)}, \quad \text{con } P(A) \text{ y } P(B) \neq 0$$

En el caso de variable continua la probabilidad de que alguna de las variables tome un valor específico, no tiene sentido:

$$P(X = x) = 0 \quad P(Y = y) = 0$$

Tiene sentido que las variables tomen valores en un intervalo tan pequeño como se quiera, por ejemplo:

$$P(X \leq x/y \leq Y \leq y + \Delta y) = F(X/Y)$$

$$\frac{P(X \leq x; y \leq Y \leq y + \Delta Y)}{P(y \leq Y \leq y + \Delta y)}$$

$$\frac{\int_{-\infty}^x \int_y^{y+\Delta y} f(l, m) dl dm}{\int_{-\infty}^{\infty} \int_y^{y+\Delta y} f(x, m) dx dm} \quad \text{si } \Delta y \rightarrow 0 \text{ luego}$$

$$\frac{\int_{-\infty}^x f(l, y) dl}{\int_y^{y+\Delta Y} f(y) dy} = f(y)$$

Puede demostrarse que la anterior es función de densidad si se integra en todo el recorrido y se obtiene 1:

$$\int_{-\infty}^{\infty} f(X/Y)dx = \int_{-\infty}^{\infty} \frac{f(x, y)}{f(y)} dx = \frac{1}{f(y)} \int_{-\infty}^{\infty} f(x, y)dx = \frac{f(y)}{f(y)} = 1$$

Función de distribución condicionada de la v.a. continua X dado Y=y₀:

$$F(x/y) = \int_{-\infty}^x f(x/y)dx = \frac{\int_{-\infty}^x f(x, y)dx}{f(y)} \quad (2.24)$$

Función de distribución condicionada de la v.a. continua Y dado X=x₀:

$$F(y/x) = \int_{-\infty}^y f(y/x)dy = \frac{\int_{-\infty}^y f(x, y)dy}{f(x)} \quad (2.25)$$

Independencia de variables aleatorias

Cuando se tienen dos variables aleatorias distribuidas conjuntamente, puede interesar saber si las variables son dependientes o independientes, ya que puede tener importancia para determinar si es correcto predecir el valor de una basándose en el valor de la otra. Recordar también aquí el concepto de independencia de sucesos:

$$P(AB) = P(A).P(B)$$

$$P(X \leq x; Y \leq y) = P(X \leq x).P(Y \leq y)$$

$$F(x, y) = F(x).F(y)$$

Esto sólo es posible de expresar si las variables son independientes. Las variables aleatorias son independientes *si y sólo si* la función conjunta puede expresarse como el producto de sus marginales. Si la variable bidimensional es discreta se expresa de la siguiente forma:

$$P(X = x_i; Y = y_j) = P(X = x_i) \cdot P(Y = y_j)$$

$$f(x_i, y_j) = f(x_i) \cdot f(y_j)$$

Por lo tanto puede concluirse que: **es necesario para que dos variables aleatorias distribuidas conjuntamente sean independientes, que su función conjunta pueda descomponerse en el producto de sus marginales.**

Ahora si la función conjunta está expresada como el producto de sus marginales, es suficiente para asegurar que las variables son independientes.

Ejemplo de variable aleatoria bidimensional

Variable continua

Si X e Y representan las duraciones, en años, de dos componentes en un sistema electrónico y la función de densidad conjunta de estas variables es la que se presenta, interesa saber si son variables dependientes o no.

$$f(x,y) = \begin{cases} e^{-(x+y)} & x > 0, y > 0 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

se debería verificar que es una función de densidad conjunta:

$$\int_0^{\infty} \int_0^{\infty} e^{-(x+y)} dy dx = 1$$

luego las funciones marginales:

$$f(x) = \int_0^{\infty} e^{-(x+y)} dy$$

$$f(y) = \int_0^{\infty} e^{-(x+y)} dx$$

para luego verificar si su producto es igual a la función conjunta.



Universidad Nacional del Litoral
Facultad de Ingeniería y Ciencias Hídricas

ESTADÍSTICA

Ingeniería Informática

TEORÍA

Mg. Ing. Susana Vanlesberg
Profesor Titular

UNIDAD 3

Características de variables aleatorias

CARACTERÍSTICAS

VARIABLE ALEATORIA UNIDIMENSIONAL

Como ya se ha expresado, no es posible predecir con total seguridad el valor futuro de una variable aleatoria, pero sí es posible realizar una descripción completa de su comportamiento a través de sus leyes de probabilidad (función de cuantía, función de densidad, función de distribución).

Existen una serie de números que resumen las características dominantes del comportamiento de una variable aleatoria. Estos números son una forma conveniente de cuantificar la ubicación y la forma de una distribución de probabilidades. Se los puede clasificar de la siguiente manera:

a - Medidas de la tendencia central

b - Medidas de variabilidad

c - Medidas de asimetría

d - Medidas de curtosis

3.1 - Medidas de tendencia central

Se distinguen, dentro de esta clase, las que son *promedios* y las que son *medidas de ubicación*.

Promedios

a- Esperanza matemática

b- Media geométrica

c- Media armónica

Medidas de ubicación

a- Mediana

b- Modo

c- Cuantiles

Promedios

a- Esperanza Matemática

Suele ser denominada también *valor medio* ó *valor esperado*. Es uno de los conceptos más importantes y es muy utilizado en la teoría de decisiones, en el análisis de sistemas y en muchos otros campos.

Se la define de la siguiente forma:

$$E(X) = \sum_{\forall i} x_i f(x_i) \quad \text{variable discreta}$$

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx \quad \text{variable continua} \quad (1)$$

En ella se condensa la información que hay en la función de probabilidad en un solo número. Geométricamente define el centro de gravedad de la masa de la distribución de probabilidades de la variable aleatoria.

Propiedades

1- La esperanza de una constante es igual a la misma constante:

$$E(C) = C$$

$$E(C) = \int_{-\infty}^{\infty} C \cdot f(x) dx = C \cdot \int_{-\infty}^{\infty} f(x) dx = C$$

2- La esperanza de una constante por una función es igual a la constante por la esperanza de la función:

$$E(C X) = C E(X)$$

$$E(CX) = \int_{-\infty}^{\infty} C X f(X) dX = C \int_{-\infty}^{\infty} X f(X) dX = C E(X)$$

3- Esperanza de una función de la variable X:

$$E(a + bX) = a + b E(X)$$

4- Esperanza de una suma de funciones de X:

$$E(g_1(x) \pm g_2(x)) = E(g_1(x)) \pm E(g_2(x))$$

5- Esperanza de un producto de variables aleatorias independientes:

$$E(X.Y) = E(X) \cdot E(Y)$$

b - Media geométrica

$$M_g = \left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}} \quad (2)$$

El logaritmo de la media geométrica es igual al valor esperado de los logaritmos de los x_i :

$$\log M_g = \sum_{\forall i} f(x_i) \log x_i \quad \text{para variable discreta} \quad (3)$$

$$\log M_g = \int_{-\infty}^{\infty} \log x f(x) dx \quad \text{para variable continua}$$

c- Media armónica

$$M_H = \frac{1}{\sum_{\forall i} \frac{1}{x_i} f(x_i)} \quad \text{para variable discreta} \quad (4)$$

$$M_H = \frac{1}{\int_{-\infty}^{\infty} \frac{1}{x} f(x) dx} \quad \text{para variable continua}$$

Medidas de ubicación

a- Mediana

Al igual que los cuantiles y el modo es una medida de ubicación o posición de la función de probabilidades.

Es el valor de la variable que cumple con la siguiente condición:

$$\int_{-\infty}^{\text{Mediana}} f(x)dx = 0.5$$

$$P(X \leq \text{Mediana}) = 0.5$$

si X es variable continua

$$\text{o bien } \int_{\text{Mediana}}^{\infty} f(x)dx = 0.5$$

$$P(X > \text{Mediana}) = 0.5$$

$$\text{Prob}(X \leq \text{Mediana}) = \sum_{i=1}^{\text{Mediana}} f(x_i) = 0.5$$

si X es variable discreta

$$\text{o bien } \text{Prob}(X > \text{Mediana}) = \sum_{\text{Mediana}}^{\infty} f(x_i) = 0.5$$

(5)

b – Modo

El modo es el valor de la variable que ocurre más frecuentemente. De esta manera, si la variable aleatoria es continua será el valor de X que maximice a la función de densidad:

$$\frac{df(x)}{dx} = 0 \quad \text{y} \quad \frac{d^2 f(x)}{dx^2} < 0 \quad (6)$$

y, si la variable es discreta, será el valor de la variable asociado a la máxima probabilidad:

$$M_x = \underset{i=1}{\overset{n}{\text{argmax}}} f(x_i) \quad (7)$$

c - Cuantiles

Se denomina *cuantil de orden p* (siendo p un número perteneciente al intervalo $[0,1]$) al valor de la variable X_p que cumple con la siguiente condición:

$$P(X \leq x_p) = p \text{ o bien } P(X > x_p) = 1 - p \quad (8)$$

Cuartiles. Existen tres cuartiles, y dividen a la distribución de probabilidades en cuatro partes; de allí su nombre. Su expresión es la siguiente:

$$P(X \leq x_{p_i}) = \frac{i}{4}, \text{ para } i = 1, 2, 3$$

El cuartil de orden 2 corresponde a la mediana.

Deciles. Dividen la distribución de probabilidades en diez partes, por lo tanto hay 9 deciles:

$$P(X \leq x_{p_i}) = \frac{i}{10}, \text{ para } i = 1, 2, \dots, 9$$

Porcentiles: Dividen la distribución de probabilidades en 100 partes; por lo tanto, hay 99 porcentiles. Son de interés cuando se desea analizar detalladamente la distribución de probabilidades:

$$P(X \leq x_{p_i}) = \frac{i}{100}, \text{ para } i = 1, 2, \dots, 99$$

3.2 - Medidas de variabilidad

Indican la dispersión o aleatoriedad en el comportamiento de la variable aleatoria.

- a - Rango.
- b - Varianza - Desvío.
- c - Coeficiente de variabilidad.

a – Rango

Es simplemente la diferencia entre el mayor y el menor valor de la variable; es una cantidad que no aporta mucha información, ya que solamente se considera un par de números. Suele variar desde $-\infty$ a $+\infty$ ó de 0 a $+\infty$, o bien entre dos números.

b - Varianza

Es la medida de dispersión más usada. Previo a definirla, se hace necesario definir una herramienta muy importante: los *momentos*.

Momentos: se los define como los promedios de distintas potencias de la variable aleatoria:

$$\alpha_k = E[X^k] = \sum_{\forall i} x_i^k \cdot f(x_i) \quad \text{ó}$$

$$\alpha_k = E[X^k] = \int_{-\infty}^{+\infty} x^k f(x) dx$$

(9)

Estas expresiones corresponden a los momentos k-ésimos del área de la función de probabilidades con respecto al origen. Si k=1, se obtiene la esperanza.

Es posible definir también los momentos de áreas con respecto a otro punto que no sea el origen. En particular, momentos con respecto a la media se denominan **momentos centrados**.

$$\mu_k = E[(X - E[X])^k] = \int_{-\infty}^{+\infty} (x - E[X])^k f(x) dx, \text{ si } X \text{ es continua}$$

(10)

$$\mu_k = E[(X - E[X])^k] = \sum_{\forall i} (x_i - E[X])^k \cdot f(x_i), \text{ si } X \text{ es discreta}$$

El momento de orden 1 es igual a 0:

$$\mu_1 = E[x - E[X]] = E[X] - E[E[X]] = E[X] - E[X] = 0$$

El momento centrado de orden 2 define a la **varianza de X**:

$$\text{Var}[X] = \sigma_x^2 = E[(X - E[X])^2]$$

$$\text{Var}[X] = \int_{-\infty}^{+\infty} (x - E[X])^2 f(x) dx, \text{ si } X \text{ es continua}$$

(11)

$$\text{Var}[X] = \sum_{\forall i} (x_i - E[X])^2 f(x_i), \text{ si } X \text{ es discreta}$$

Esta expresión puede simplificarse y ser expresada en función de los momentos con respecto al origen, por desarrollar la potencia del binomio:

$$\begin{aligned}
E[X - E[X]]^2 &= E[X^2 - 2XE[X] + E^2[X]] = \\
&= E[X^2] - 2E[X]E[X] + E[E^2[X]] = \\
&= E[X^2] - 2E^2[X] + E^2[X] = \\
&= E[X^2] - E^2[X] = \alpha_2 - \alpha_1^2
\end{aligned}
\tag{12}$$

Propiedades

1- La varianza de una constante es igual a cero:

$$\text{Var}(C) = 0$$

$$E[C - E(C)]^2 = E[C - C]^2 = E(0) = 0$$

2- La varianza de una constante por X es igual a la constante al cuadrado por la varianza de la variable X:

$$\text{Var}(CX) = C^2 \text{Var}(X)$$

$$\begin{aligned}
E[CX - E(CX)]^2 &= E[CX - CE(X)]^2 = \\
&= E\{[C.(X - E(X))]^2\} = E\{C^2.[X - E(X)]^2\} = \\
&= C^2 E[X - E(X)]^2 = C^2 \text{Var}(X)
\end{aligned}$$

3- Varianza de una función de X:

$$\text{Var}(a + bX) = \text{Var}(a) + \text{Var}(bx) = 0 + b^2 \text{Var}(X)$$

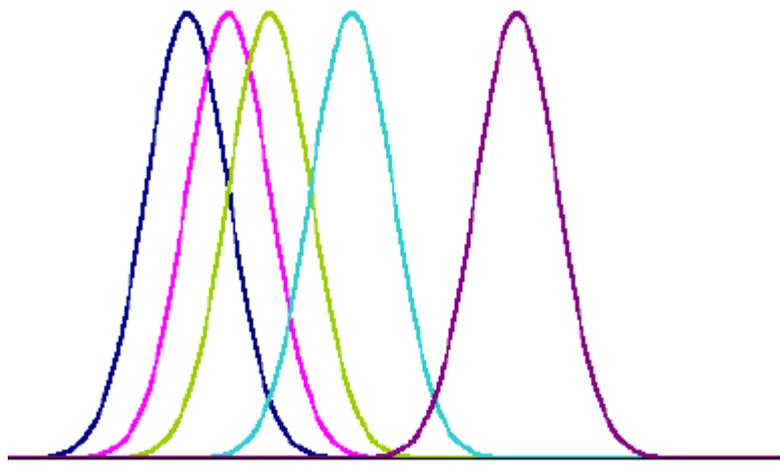
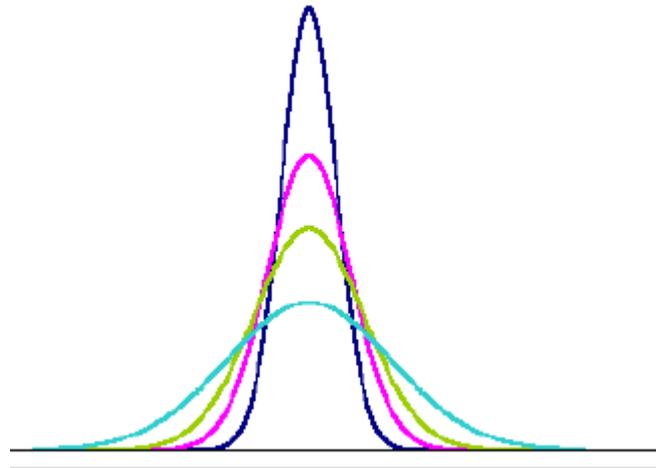
Desvío Estándar

Para algunas aplicaciones suele ser más conveniente utilizar una medida de variabilidad en las mismas unidades de la variable; para eso se define el *desvío estándar* como la raíz cuadrada positiva de la varianza:

$$\sigma_x = +\sqrt{\text{Var}(X)} \quad (13)$$

$$\sigma_x = +\sqrt{\alpha_2 - \alpha_1^2}$$

En las siguientes figuras se observa el efecto de variar μ y σ .



Coefficiente de Variabilidad

Es un coeficiente adimensional que se obtiene de dividir el desvío por el valor esperado. Se lo utiliza para comparar las dispersiones de poblaciones que correspondan a variables con diferentes unidades de medida o bien en aquellos casos en que si bien las unidades son iguales los valores medios sean muy diferentes:

$$C_v = \frac{\sigma(X)}{E(X)} \quad (14)$$

3.3 - Medidas de asimetría

De igual forma que la media y la varianza miden la ubicación y dispersión de una distribución, los momentos más altos miden otras propiedades de la misma.

El tercer momento respecto a la media es usado para determinar si una distribución es simétrica o asimétrica.

Si la *distribución es simétrica*, como las desviaciones están elevadas al cubo, las positivas y negativas tienden a anularse y, por lo tanto, $\mu_3 = 0$.

Si la *distribución es asimétrica a la derecha*, $\mu_3 < 0$.

Si la *distribución es asimétrica a la izquierda*, $\mu_3 > 0$.

μ_3 , tomado como valor aislado, no es una buena medida de la asimetría, ya que tiene las mismas unidades que la variable; por eso es que se define una medida relativa denominada *coeficiente de asimetría*:

$$\gamma_1 = \frac{\mu_3}{\sigma^3} \quad (15)$$

Esto puede expresarse en función de los momentos respecto al origen, desarrollando μ_3 :

$$\begin{aligned} \mu_3 &= E(X - E[X])^3 = E[X^3 - 3X^2E[X] + 3XE^2[X] - E^3[X]] = \\ &= E[X^3] - 3E[X^2]E[X] + 3E^3[X] - E^3[X] = \\ &= \alpha_3 - 3\alpha_2\alpha_1 + 3\alpha_1^3 - \alpha_1^3 \end{aligned}$$

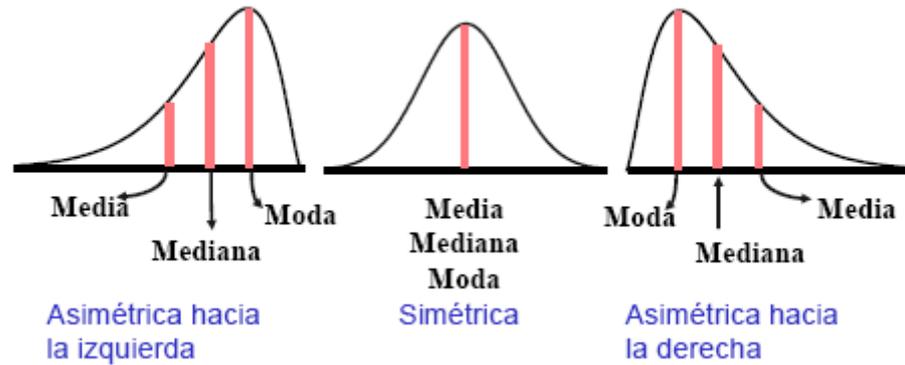
$$\mu_3 = \alpha_3 - 3\alpha_2\alpha_1 + 2\alpha_1^3$$

$$\gamma_1 = \frac{\alpha_3 - 3\alpha_2\alpha_1 + 2\alpha_1^3}{(\alpha_2 - \alpha_1^2)^{3/2}}$$

Se cumplen también las siguientes relaciones aproximadas:

$$As = \frac{E(X) - \text{Modo}}{\sigma(X)} \quad As = \frac{3 [E(X) - \text{Mediana}]}{\sigma(X)}$$

La distribución será simétrica si $As=0$; asimétrica a la derecha si $As<0$; y asimétrica a la izquierda si $As>0$.



3.4 - Medidas de curtosis

Una cuarta propiedad de las variables aleatorias se basa en el momento centrado de cuarto orden, que permite evaluar el empinamiento o aplastamiento de la distribución de probabilidades comparada con una curva tomada como modelo (la curva normal).

Puede introducirse una medida relativa para independizarse de las unidades:

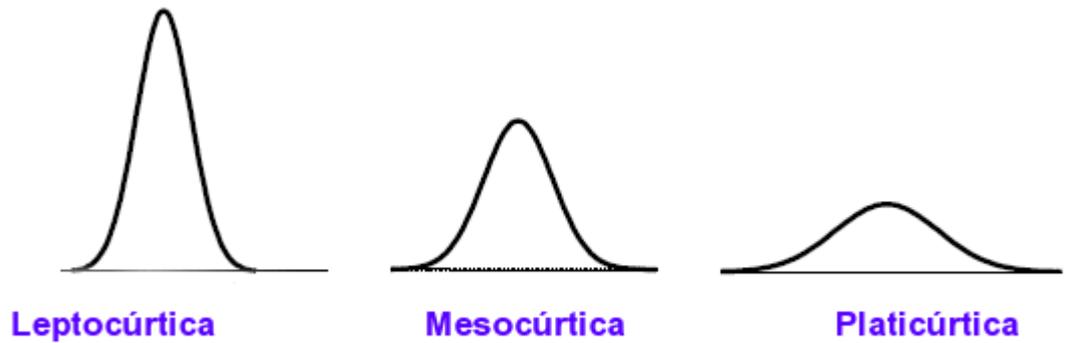
$$\gamma_2 = \frac{\mu_4}{\sigma^4} \quad (16)$$

Desarrollando μ_4 en función de los momentos respecto al origen:

$$\begin{aligned} \mu_4 &= E[X - E[X]]^4 = E[X^4 - 4X^3 E[X] + 6X^2 E^2[X] - 4X E^3[X] + E^4[X]] = \\ &= E[X^4] - 4E[X^3]E[X] + 6E[X^2]E^2[X] - 4E[X]E^3[X] + E^4[X] \end{aligned}$$

$$\mu_4 = \alpha_4 - 4\alpha_3\alpha_1 + 6\alpha_2\alpha_1^2 - 3\alpha_1^4$$

La Curtosis de la curva Normal es igual a 3, y se la denomina distribución *mesocúrtica*. Para distribuciones que presenten mayor concentración de probabilidad cerca de la media, mayor que en la Normal, la curtosis será mayor que 3 y se denominará *leptocúrtica*. En caso que la concentración alrededor de la media sea menor que en la Normal la curtosis será menor que 3 y la distribución se dice *platicúrtica*.



3.5 - Momentos de variables aleatorias distribuidas conjuntamente

Se parte de considerar una función de las variables X e Y de la siguiente forma:

$$g(X,Y) = X^l Y^n$$

Para obtener el momento conjunto de orden l,n, se obtiene la esperanza de la expresión anterior, desarrollada para el caso de una variable bidimensional continua:

$$E(X^l, Y^n) = \sum_{\forall x_i} \sum_{\forall y_j} x^l y^n f(x, y) \quad (17)$$

Los momentos más importantes son los de orden (1,0), (0,1), (2,0) y (0,2).

En el caso que $g(X,Y) = X$, se obtiene la esperanza de X:

$$E[g(X,Y)] = E(X) = \sum_{\forall x_i} \sum_{\forall y_j} x_i f(x_i, y_j)$$

$$E(X) = \alpha_{1,0} = \sum_{\forall x_i} x_i \sum_{\forall y_j} f(x_i, y_j)$$

$$E(X) = \alpha_{1,0} = \sum_{\forall x_i} x_i f(x_i) \quad (18)$$

Esta expresión es similar a la obtenida para el caso de una variable; de esta manera la esperanza de X es el valor esperado de X sin considerar a Y.

Para obtener la esperanza de Y se procede de forma análoga:

$$E(Y) = \sum_{\forall x_i} \sum_{\forall y_j} y_j f(x_i, y_j) = \sum_{\forall y_j} y_j \sum_{\forall x_i} f(x_i, y_j)$$

$$E(Y) = \alpha_{0,1} = \sum_{\forall y_j} y_j f(y_j) \quad (19)$$

$\alpha_{1,0}$ y $\alpha_{0,1}$ establecen el centro de masa de la distribución de probabilidades.

3.5.1 - Momentos centrados

Como en el caso de una variable, y como en mecánica, el momento centrado más usado es el de segundo orden. Tomando ahora una función de ambas variables igual a:

$$g(X,Y) = [X - E(X)]^l \cdot [Y - E(Y)]^n$$

y hallando la esperanza de esta función, se obtiene :

$$E[g(x,y)] = E \{ [x - E(X)]^l \cdot [y - E(Y)]^n \} \quad (20)$$

De la misma manera que para el caso unidimensional los momentos centrados de orden 1,0 o 0,1 son iguales a cero.

Los momentos más usados son los de orden 2,0, 0,2 y 1,1. Los momentos de orden 2,0 y 0,2 llevan a obtener las varianzas marginales.

Expresadas en función de los momentos respecto al origen,

$$\sigma^2(X) = \alpha_{2,0} - \alpha_{1,0}^2 \quad (21)$$

Un resultado similar se obtiene para la varianza de Y:

$$\sigma^2(Y) = \alpha_{0,2} - \alpha_{0,1}^2 \quad (22)$$

Covarianza

Un momento muy importante se obtiene de considerar l=1 y n=1. Este momento se denomina *covarianza*.

$$\begin{aligned} Cov(X,Y) = \sigma_{x,y} &= E \{ [X - E(X)] [y - E(Y)] \} = \\ &= \sum_{\forall x_i} \sum_{\forall y_j} [x_i - E(X)] [y_j - E(Y)] f(x_i, y_j) \end{aligned} \quad (23)$$

Puede ser expresada en función de los momentos respecto al origen por desarrollar la ecuación anterior:

$$\sigma_{x,y} = \alpha_{1,1} - \alpha_{1,0} \cdot \alpha_{0,1}$$

Coefficiente de correlación

Una versión normalizada de la covarianza es el denominado *coeficiente de correlación* $\rho(x,y)$, y se obtiene dividiendo la covarianza por el producto de sus desviaciones:

$$\begin{aligned} \rho_{x,y} &= \frac{COV_{x,y}}{\sigma_x \cdot \sigma_y} = \frac{\mu_{1,1}}{\sqrt{\mu_{2,0} \cdot \mu_{0,2}}} \\ &-1 \leq \rho \leq 1 \end{aligned} \quad (24)$$

Este coeficiente merece algunos comentarios:

-de la misma manera que la media y la varianza, el coeficiente de correlación es útil cuando la ley de probabilidades no se conoce en forma completa; en este caso pares de observaciones de ambas variables pueden ayudar a estimar el coeficiente de correlación y poder obtener alguna conclusión acerca de su comportamiento a través de su valor.

-de la definición de covarianza se ve que valores positivos de ella resultarán de pares de valores altos de x con valores altos de y o valores pequeños de x y pequeños de y; mientras que valores negativos de la covarianza se darán con la asociación de valores pequeños de una con valores grandes de la otra. En ambos casos es posible decir que existe al menos alguna dependencia estocástica entre ambas.

-si dos variables aleatorias son independientes, su covarianza y por lo tanto su coeficiente de correlación serán iguales a cero. La independencia implica que $f(x,y)$ puede descomponerse en el producto de las marginales $f_1(x) \cdot f_2(y)$ por lo tanto la covarianza se puede obtener como:

$$\begin{aligned} & \sum_{\forall x_i} \sum_{\forall y_j} [x_i - E(X)] \cdot [y_j - E(Y)] \cdot f(x) \cdot f(y) = \\ & = \sum_{\forall x_i} [x_i - E(X)] \cdot f(x_i) \sum_{\forall y_j} [y_j - E(Y)] \cdot f(y_j) = \\ & = E[x_i - E(X)] \cdot E[y_j - E(Y)] = \mu_{1,0} \cdot \mu_{0,1} \end{aligned}$$

siendo $\mu_{1,0} = 0$ y $\mu_{0,1} = 0$

-lo inverso no es cierto, o sea si el coeficiente de correlación o la covarianza son cero esto no implica que las variables sean independientes, puede haber alta dependencia estocástica y ser $\rho=0$. **Más específicamente, el coeficiente de correlación es una medida de dependencia lineal entre dos variables aleatorias.**



Universidad Nacional del Litoral
Facultad de Ingeniería y Ciencias Hídricas

ESTADÍSTICA

Ingenierías: Recursos Hídricos-Ambiental-Agrimensura

TEORÍA

Mg.Ing. Susana Vanlesberg
Profesor Titular

UNIDAD 4

Modelos Probabilísticos

Variable Discreta

La aplicación de la teoría de probabilidad en situaciones reales, concretas, que presentan una regularidad en el desarrollo, ha originado una serie de modelos que permiten resolver esas situaciones. Los ingenieros realizan supuestos respecto del problema a resolver, esto los lleva a descripciones análogas y a formas matemáticamente iguales a las de los modelos probabilísticos.

En este capítulo no sólo se presentan los modelos más usados en la práctica, sino que también se brindan algunas nociones de los mecanismos por los cuales se ha originado cada distribución. Esto último es de suma importancia para un ingeniero, ya que la existencia de tales mecanismos pueden describir una situación física de su interés, y ésta es una razón más importante que la buena aproximación matemática de algún modelo.

MODELOS DE VARIABLE ALEATORIA DISCRETA

Cuando en el desarrollo de los modelos se mencionen pruebas, éstas se referirán a observaciones sistemáticas de algún fenómeno natural.

Modelo Bernoulli

Tal vez la situación más común que se presenta es aquella en la que los resultados de los experimentos pueden separarse en dos categorías mutuamente excluyentes: *éxito* o *fracaso*; por ejemplo llueve o no, mido o no mido, está contaminado o no, etc.

Puede entonces definirse la variable aleatoria x de tipo Bernoulli y asignársele valores (arbitrarios pero muy prácticos) a los eventos antes mencionados:

$x=0$ fracaso

$x=1$ éxito

La función de probabilidad de x es simplemente:

$$f(x) = \begin{cases} p, & x = 1 \\ 1 - p, & x = 0 \end{cases} \quad (1)$$

p probabilidad de éxito.

Esta función deberá cumplir con las condiciones de una función de probabilidad:

$$\sum_{x_i=0}^{x_i=1} f(x_i) = 1$$

$$p + (1 - p) = p + 1 - p = 1$$

Las características de este modelo:

$$E(x) = \sum_{i=0}^1 x_i f(x_i) = 0(1 - p) + 1p = p \quad (2)$$

$$V(x) = \sum_{-x_i} (x_i - E(x))^2 f(x_i) =$$

$$= (0 - p)^2(1 - p) + (1 - p)^2 \cdot p =$$

$$= p^2(1 - p) + (1 - p)^2 \cdot p =$$

$$= p(1 - p)(p + 1 - p)$$

$$V(x) = p(1 - p) \quad (3)$$

Modelo Binomial

Se realizan una serie de pruebas de tipo Bernoulli cuyos resultados son mutuamente independientes y si la probabilidad de éxito permanece invariable en todas ellas se origina un nuevo modelo denominado **BINOMIAL**.

Para determinar la distribución correspondiente se analiza el número total de éxitos en n pruebas Bernoulli cada una con probabilidad favorable igual a p .

Se consideran por ejemplo 3 pruebas y se analizarán las probabilidades de éxito:

- Ningún éxito: $x=0$

0 ; 0 ; 0 ;

$(1 - p)^3$

- Un éxito: $x=1$

1 ; 0 ; 0 ; ó 0 ; 1 ; 0 ; ó 0 ; 0 ; 1

cada secuencia es un evento, que entre sí son mutuamente excluyentes cuya probabilidad es :

$p(1-p)^2$; de esta manera la probabilidad de un éxito es :

$$3p(1-p)^2$$

- Dos éxitos: $x=2$

1; 1; 0; ó 1; 0; 1; ó 0; 1; 1;

nuevamente cada secuencia es un evento que es excluyente de los restantes:

$p^2(1-p)$ y el total :

$$3p^2(1-p)$$

- Tres éxitos: $x=3$

1; 1; 1; p; p; p

$$p^3$$

Esto permite generalizar y obtener la función de probabilidad de este modelo:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

(4)

$$\text{coeficiente binomial} \binom{n}{x} = \frac{n!}{x!(n-x)!}$$

El coeficiente binomial tiene en cuenta el número de formas en que exactamente x éxitos se presenten en n pruebas. Los parámetros de este modelo son n y p ; n entero y p un número comprendido entre 0 y 1.

El modelo se denomina **BINOMIAL** porque puede considerarse como el desarrollo del binomio $(p + q)^n$.

La función debe cumplir las condiciones de una función de probabilidad:

$$\begin{aligned}
 \sum_{x_i} f(x_i) &= 1 \\
 \sum_{x_i} \binom{n}{x_i} p^{x_i} q^{n-x_i} &= \\
 &= \sum_{x_i=0}^n \frac{n!}{x_i!(n-x_i)!} p^{x_i} q^{n-x_i} = \\
 &= \frac{n!}{0!n} p^0 q^n + \frac{n!}{1!(n-1)!} p q^{n-1} + \dots + \frac{n!}{n!0!} p^n q^0 = \\
 &= q^n + \frac{n(n-1)!}{(n-1)!} p q^{n-1} + \dots + p^n = (p+q)^n = 1^n = 1
 \end{aligned}$$

La función de distribución puede obtenerse aplicando la definición:

$$F(x) = P(X \leq x) = \sum_{\forall x_i \leq x} f(x_i) \quad (5)$$

Características del modelo

$$\begin{aligned}
 E(X) &= \sum_{x_i=0}^n x_i f(x_i) = \\
 &= 0 \cdot \frac{n!}{0!(n-0)!} p^0 q^n + 1 \cdot \frac{n!}{1!(n-1)!} p^1 q^{n-1} + \dots + n \cdot \frac{n!}{n!(n-n)!} p^n q^0 = \\
 &= 0 + \frac{n(n-1)!}{(n-1)!} p q^{n-1} + \dots + n p^n q^0 = np(q^{n-1} + \dots + (n-1)p^{n-2}q + p^{n-1}) = \\
 &= np(p+q)^{n-1}
 \end{aligned}$$

$$E(x) = np \quad (6)$$

$$\text{Var}(X) = E[X - E(X)]^2 = E(X^2) - E^2(X)$$

$$E(X^2) = \sum_{x_i=0}^n \binom{n}{x_i} p^{x_i} q^{n-x_i} x_i^2 =$$

ya que el primer termino es cero
la suma se toma desde 1

$$\begin{aligned} &= np \sum_{x_i=1}^n \frac{(n-1)!}{x_i!(n-x_i)!} p^{x_i-1} q^{n-x_i} x_i^2 \\ &= np \sum_{x_i=1}^n \frac{(n-1)!}{x_i(x_i-1)!(n-x_i)!} p^{x_i-1} q^{n-x_i} x_i^2 \end{aligned}$$

haciendo $x_i = x_j + 1$, para $x_i = 1, x_j = 0, x_i = n, x_j = n-1$
 $x_i! = (x_i - 1)! x_i = x_j!(x_j + 1)$

$$\begin{aligned} &np \sum_{x_j=0}^{n-1} \frac{(n-1)!}{(x_j)!(x_j+1)(n-x_j-1)!} p^{x_j} q^{n-1-x_j} (x_j+1) = \\ &np \left[\sum_{x_j=1}^{n-1} \binom{n-1}{x_j} p^{x_j} q^{n-1-x_j} x_j + \sum_{x_j=0}^{n-1} \binom{n-1}{x_j} p^{x_j} q^{n-1-x_j} \right] \end{aligned}$$

El 2do termino de la suma es 1; se analiza el 1º,

$$\text{ya que } \binom{n-1}{x_j} x_j = \frac{(n-1)! x_j}{x_j!(n-1-x_j)!} = \frac{(n-2)!(n-1)x_j}{(x_j-1)! x_j(n-2-(x_j-1))!}$$

sacando factor comun $(n-1)p$, se consigue

$$\sum_{x_j=1}^{n-1} \binom{n-2}{x_j-1} p^{x_j-1} q^{n-1-x_j}$$

haciendo $x_j = x_k + 1$, x_k varía de 0 a $n-2$ con lo cual

$$\sum_{x_k=0}^{n-2} \binom{n-2}{x_k} p^{x_k} q^{n-2-x_k} = 1$$

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

$$\text{Var}(X) = npq \quad (7)$$

Estos mismos resultados podrían haberse obtenido de forma mucho más sencilla por considerar a la variable aleatoria X como la suma de n variables independientes idénticamente distribuidas como Bernoulli, cuya esperanza es p y

$$E(X) = E(x_1 + x_2 + \dots + x_n) = \sum_{i=1}^n E(x_i) = n E(x_i) = n p$$

varianza $p \cdot q$:

$$Var(X) = \sum_{i=1}^n Var(x_i) = n p(1 - p) = n p q$$

La distribución Binomial ha sido tabulada, tanto su función de cuantía como la de distribución. Para el caso de la función masa existen valores hasta $p = 0.5$. En caso de necesitarse $p > 0.5$, se debería plantear el problema cambiando el suceso en estudio por su contrario y tener cuidado en la interpretación de los resultados. También en los software que poseen fórmulas estadísticas y en los específicos se pueden encontrar los valores de esta distribución para cualquier valor de p .

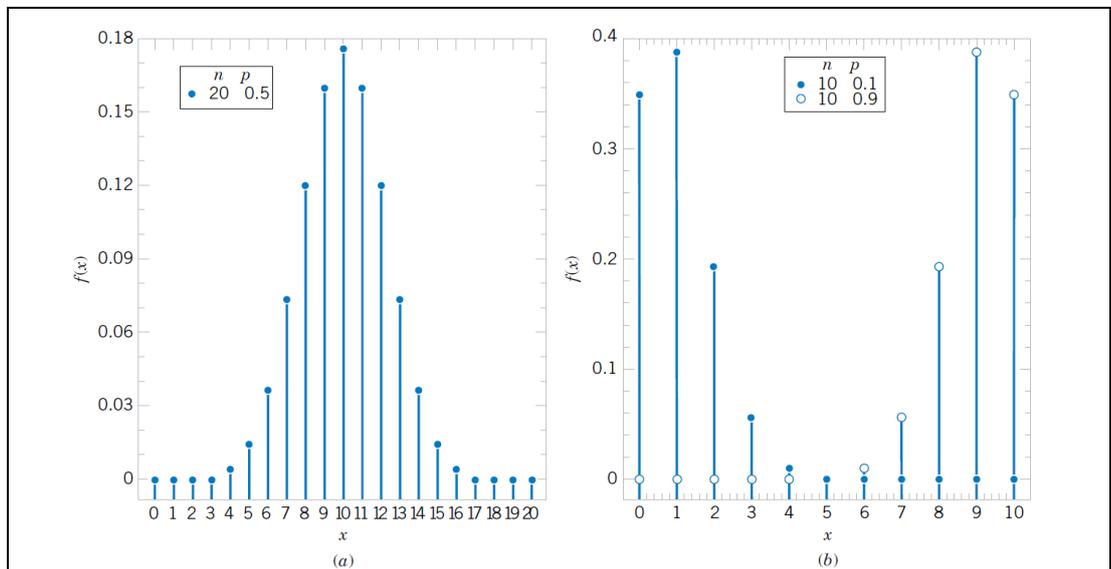


Figura N° 1 - Distribución Binomial para distintos valores de n y p

Ejemplo

Cada muestra de agua tiene una probabilidad del 10% de contener un contaminante orgánico particular. Asumir que las muestras son independientes con respecto a la presencia del contaminante. Encontrar la probabilidad de que en las próximas 18 muestras, exactamente 2 contengan el contaminante y el valor medio de muestras contaminadas.

Solución:

La variable será:

$X =$ " número de muestras que contienen el contaminante"

$$X \sim B(n=18; p=0,1)$$

$$P(X = 2) = \binom{18}{2} 0.1^2 0.9^{16} = 0.284$$

$$E(X) = np = 18 \cdot 0.1 = 1.8 \cong 2$$

Modelo Geométrico

Continuando con el mismo esquema de pruebas repetidas de tipo Bernoulli, puede interesar conocer en que prueba ocurrirá el primer éxito.

Las pruebas son repetidas, independientes, con dos posibles resultados: ocurrencia o no del evento en estudio, con probabilidad favorable $p = \text{constante}$; se deriva el modelo correspondiente a la variable N : número de pruebas hasta que ocurre el primer éxito. Este primer éxito se producirá sí y sólo sí en las pruebas anteriores no se produjo, lo cual ocurre con probabilidad igual a $(1 - p)$. Luego su función de probabilidad es:

$$P(N = n) = f(n) = (1 - p)^{n-1} p \quad \text{con } n = 1, 2, 3, \dots (8)$$

Este modelo se denomina **GEOMÉTRICO** de parámetro p .

La función acumulativa se obtiene a partir de la función anterior:

$$F(n) = \sum_{j=1}^n p(j) = \sum_{j=1}^n (1 - p)^{j-1} p =$$

$$= p \left[\frac{(1 - p)^n - 1}{(1 - p) - 1} \right]$$

$$\left[\frac{(1 - p)^n - 1}{(1 - p) - 1} \right]$$

*desarrollo de la
suma de una progresión geométrica*

$$F(x) = 1 - (1 - p)^n \quad (9)$$

Esto mismo se puede obtener al considerar la probabilidad de que $N \leq n$ simplemente como la probabilidad que exista al menos una ocurrencia en N pruebas:

$$P(N \geq 1) = 1 - P(N < 1) = 1 - P(\text{no haya ocurrencias en } n \text{ pruebas}) =$$

$$F(x) = 1 - (1 - p)^n$$

La gráfica de la función de cuantía:

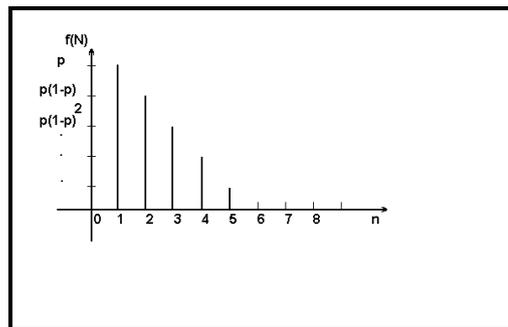


Figura N° 2-Función del modelo Geométrico

Características

$$E(N) = \sum_{n=1}^{\infty} n p (1-p)^{n-1} = p \sum_{n=1}^{\infty} n (1-p)^{n-1}$$

$$\sum_{n=1}^{\infty} n (1-p)^{n-1} = 1(1-p)^0 + 2(1-p) + 3(1-p)^2 + 4(1-p)^3 + \dots =$$

$$= 1 + 2q + 3q^2 + 4q^3 + \dots = \frac{1}{(1-q)^2}$$

$$E(N) = p \frac{1}{p^2} = \frac{1}{p} \quad (10)$$

Este resultado significa que el número promedio de pruebas hasta que ocurre el primer éxito es la inversa de la probabilidad p de ocurrencia del evento. Esto se asocia al concepto de **período de retorno**, sumamente utilizado en ingeniería.

$$\begin{aligned}
\text{Var}(N) &= E(N^2) - E^2(N) = \left[\sum_{n=1}^{\infty} n^2 p (1-p)^{n-1} \right] - \left(\frac{1}{p} \right)^2 = \\
&= \frac{1}{p^2} \left[\sum_{n=1}^{\infty} n^2 p^3 (1-p)^{n-1} - 1 \right] = \\
&= \frac{1}{p^2} \left[\sum_{n=1}^{\infty} n^2 (1-q)^3 q^{n-1} - 1 \right] = \\
&= \frac{1}{p^2} \left[\sum_{n=1}^{\infty} q^{n-1} (1-3q+3q^2+q^3) n^2 - 1 \right] \\
&\quad \sum_{n=1}^{\infty} q^{n-1} (1-3q+3q^2+q^3) n^2 = 1+q \\
\text{Var}(N) &= \frac{1}{p^2} (1+q-1)
\end{aligned}$$

$$\text{Var}(N) = \frac{1-p}{p^2} \quad (11)$$

El término período de retorno es muchas veces mal interpretado ya que se cree que pasarán 100 años entre ocurrencias de valores de una variable de interés superiores a un valor crítico; en realidad la probabilidad de que se produzca un valor mayor al crítico permanece invariable y en este caso igual a 0.01 para cualquier año, independientemente de lo que ocurrió en años anteriores, se está hablando de valor promedio.

Ejemplo

Cuál sería la probabilidad de que una crecida con período de retorno 10 años ocurra antes o durante el décimo año en la ciudad de Santa Fe.

$$\text{siendo } T = 1/p \quad \rightarrow \quad p = 1/T = 1/10$$

$$\text{luego } P(X \leq 10) = 1 - (1 - 0.1)^{10} = 65\%$$

$$F(10) = 1 - (1 - p)^{10}$$

Modelo Poisson

Considerando las condiciones que caracterizan al modelo Binomial, es necesario analizar que sucede con la distribución de la variable si el número de pruebas n se incrementa y la probabilidad p de éxito se hace cada vez más pequeña.

Si n se incrementa y p se hace pequeña, pero el número promedio de eventos en el intervalo total permanece constante e igual a λ denominando a esta constante λ y considerando la función masa de probabilidad de x en el límite, esto es con $p \rightarrow 0$ y $n \rightarrow \infty$ y $\lambda = n \cdot p$ a partir de la función del modelo Binomial:

$$\begin{aligned} \lambda = n p &\rightarrow p = \frac{\lambda}{n} \\ f(x) &= \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} = \\ &= \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \frac{n!}{(n-x)! n^x \left(1 - \frac{\lambda}{n}\right)^x} = \\ &= \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \left[\frac{n(n-1)(n-2)\dots(n-(x+1))(n-x)!}{(n-x)! \left(n \left(1 - \frac{\lambda}{n}\right)\right)^x} \right] \end{aligned}$$

La expresión entre corchetes tiene x términos en el numerador y x términos en el denominador, y para un n suficientemente grande, este término es igual a :

$$n^x / n^x = 1$$

con lo cual queda por evaluar la expresión $\left(1 - \frac{\lambda}{n}\right)^n$ que cuando $n \rightarrow \infty$ es igual a la constante $e^{-\lambda}$, esto permite expresar a la función de probabilidad del modelo de Poisson de la siguiente forma:

$$f(x) = P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad (12)$$

λ parámetro del modelo.

Esta función deberá cumplir con la condición para ser una función de probabilidad:

$$\begin{aligned} \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} &= \\ &= e^{-\lambda} \left[\frac{\lambda^0}{0!} + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots + \frac{\lambda^n}{n!} \right] \\ \left[\frac{\lambda^0}{0!} + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots + \frac{\lambda^n}{n!} \right] &= e^{\lambda} \\ e^{-\lambda} \cdot e^{\lambda} &= e^{-\lambda+\lambda} = e^0 = 1 \end{aligned}$$

Características

$$\begin{aligned} E(X) &= \sum_{x=0}^{\infty} x \frac{e^{-\lambda} \lambda^x}{x!} = \\ &= \lambda \sum_{x=1}^{\infty} x \frac{e^{-\lambda} \lambda^{x-1}}{x(x-1)!} \\ &= \sum_{x=1}^{\infty} \frac{e^{-\lambda} \lambda^{x-1}}{(x-1)!} = 1 \end{aligned}$$

$$E(X) = \lambda \quad (13)$$

$$\text{Var}(X) = E(X^2) - E^2(X)$$

$$\begin{aligned} E(X^2) &= \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda} \lambda^x}{x!} = \lambda \left[\sum_{x=0}^{\infty} \frac{x^2 \lambda^{x-1} e^{-\lambda}}{x(x-1)!} \right] = \\ &= \lambda \left[\sum_{x=1}^{\infty} \frac{x \lambda^{x-1} e^{-\lambda}}{(x-1)!} \right] \end{aligned}$$

si $y = x - 1 \rightarrow y + 1 = x$, luego

$$\lambda \left[\sum_{y=0}^{\infty} (y+1) \frac{\lambda^y e^{-\lambda}}{y!} \right] = \lambda \left[\sum_{y=0}^{\infty} y \frac{e^{-\lambda} \lambda^y}{y!} + \sum_{y=0}^{\infty} \frac{e^{-\lambda} \lambda^y}{y!} \right]$$

$$\sum_{y=0}^{\infty} y \frac{e^{-\lambda} \lambda^y}{y!} = E(Y) = \lambda$$

$$\sum_{y=0}^{\infty} \frac{e^{-\lambda} \lambda^y}{y!} = 1$$

luego

$$E(X^2) = \lambda(\lambda + 1) = \lambda^2 + \lambda$$

$$\text{Var}(X) = \lambda^2 + \lambda - \lambda^2$$

$$\text{Var}(X) = \lambda \quad (14)$$

Generalmente este modelo se vincula a aquellos eventos que ocurren en una unidad de tiempo, luego el período de tiempo en el que se realiza el análisis constituye una secuencia de pruebas independientes cada una con distribución Binomial.

Si se tomara para el análisis un intervalo de tiempo igual al doble o al triple del inicial el parámetro es también igual al doble, al triple, etc., marcando esto la dependencia del tiempo de este modelo y por ello vinculado a los procesos estocásticos. Se entiende por procesos estocásticos a aquellos en los que interesa la secuencia en el tiempo de ocurrencia de eventos.

Dado un intervalo de números reales, asumiendo que ocurren al azar en todo el intervalo valores de una variable aleatoria; si el intervalo se puede dividir en subintervalos de pequeña longitud tal que:

- (1) la probabilidad de que más de una ocurrencia en un subintervalo es cero,
- (2) la probabilidad de una ocurrencia en un subintervalo es la misma para todos los subintervalos y proporcional a la longitud del subintervalo, y
- (3) el conteo en cada subintervalo es independiente de otros subintervalos,

Entonces el experimento aleatorio se denomina un **proceso de Poisson**.

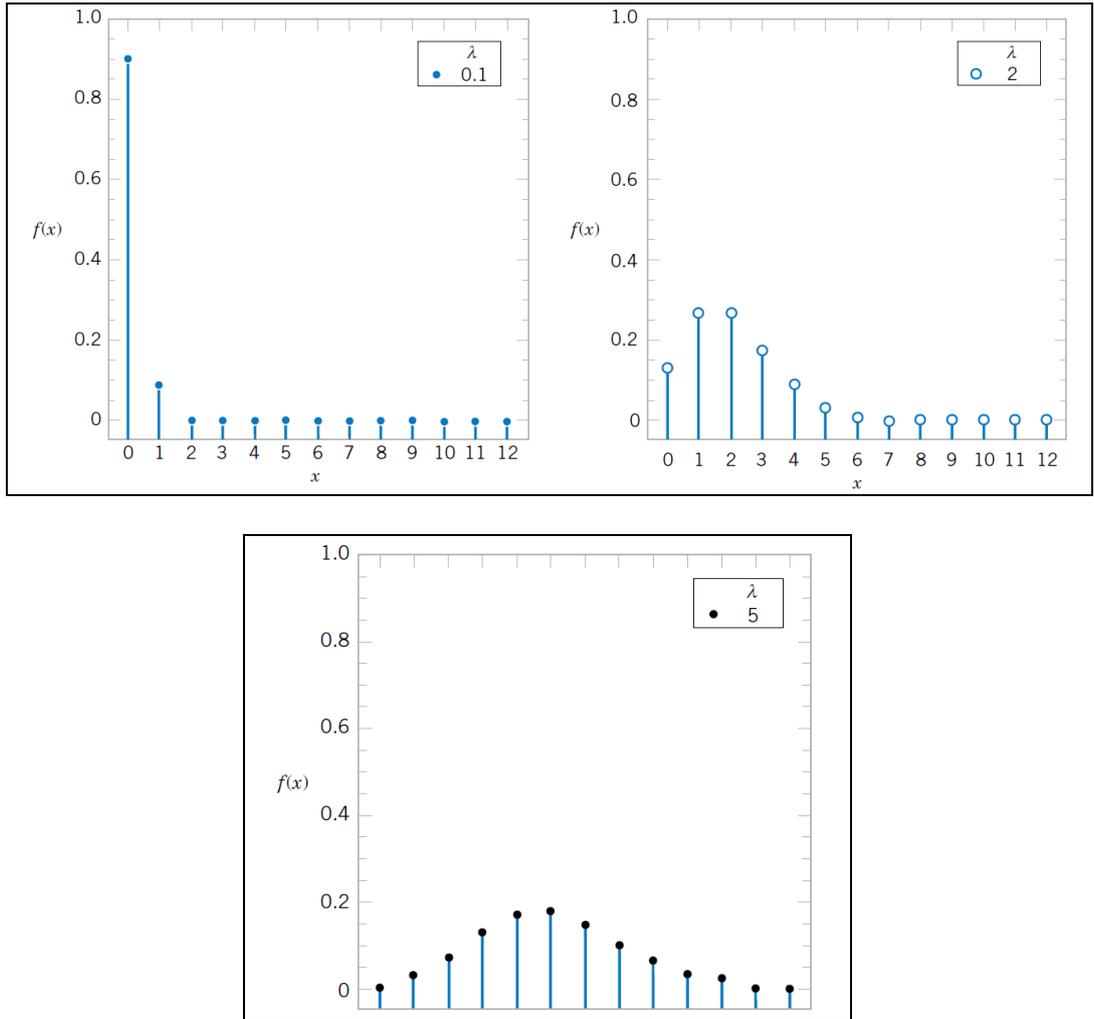


Figura N° 3 - Distribución de Poisson para distintos valores del parámetro

Ejemplo

En cierta región se sabe que en un determinado mes del año ocurren, en promedio, 3 tormentas cuya precipitación supera los 150 mm. Determine la probabilidad de que en un determinado mes en esta zona ocurran menos de 5 tormentas de este tipo.

$$\lambda = 3$$

$$P(x < 5) = \sum_{x_i=0}^4 \frac{e^{-3} 3^{x_i}}{x_i!} = 0.8153$$

Modelo Hipergeométrico

Este modelo surge cuando se realiza un muestreo sin reposición de una población finita y con sus elementos clasificados en dos categorías.

Si N es el total de elementos de los cuales hay k de una categoría y $N-k$ de otra, al realizar una extracción de n elementos, *sin reposición*, cada extracción que se realice posteriormente es dependiente del resultado de la extracción anterior con lo cual va cambiando la probabilidad de éxito.

Para derivar la función correspondiente a la variable aleatoria x : número de éxitos o elementos pertenecientes a la categoría que se estudia, en una extracción, se deberán considerar todas las maneras posibles o combinaciones, de extraer x elementos de la categoría deseada, de los n extraídos y los restantes que pertenezcan a la otra categoría. El total de casos se obtiene de las combinaciones del total N extraídos de n . Luego la función de probabilidad para este modelo es:

$$P(X = x) = f(x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} \quad (15)$$

Características

Para obtener las características es posible decir que la variable aleatoria X es la suma de n variables x_i como en el caso Binomial, pero con la diferencia que aquí las x_i son dependientes. Como para sumar las esperanzas no se necesita que las variables aleatorias sean independientes es posible obtener la esperanza de la siguiente forma:

$$E(X) = E(x_1) + E(x_2) + \dots + E(x_n)$$

donde en cada $E(x_i)$ la probabilidad de x en la i ésima prueba es k/N , si no se sabe que ha ocurrido en pruebas anteriores o posteriores, luego:

$$E(X) = n p = n \left(\frac{k}{N} \right) \quad (16)$$

La varianza:

$$Var(X) = n * p * q * \left(\frac{N-n}{N-1} \right) \quad (17)$$

siendo $\left(\frac{N-n}{N-1}\right)$ el factor de corrección por muestreo sin reposición y población finita.

Cuando $\left(\frac{n}{N}\right) \leq 0.05$ la distribución Hipergeométrica se aproxima a la Binomial.

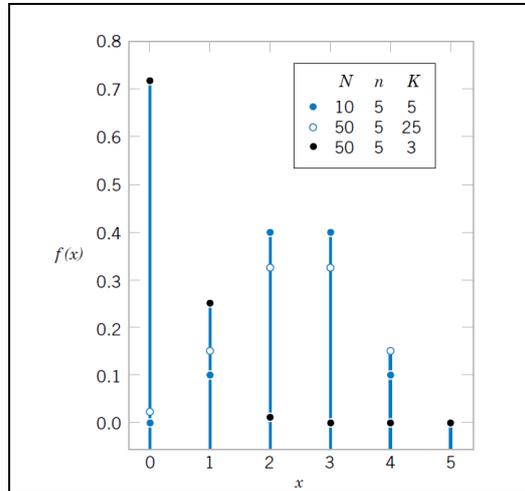


Figura N° 4 - Distribución Hipergeométrica para distintos valores de N, n y K

Ejemplo

Un equipo incluye cinco ingenieros ambientales y nueve agrimensores. Si se eligen al azar cinco profesionales y se les asigna un proyecto, ¿cuál es la probabilidad de que el equipo del proyecto incluya exactamente a dos ingenieros ambientales?.

Solución:

X: "número de ing. ambientales incluidos en el proyecto".

$$P(\text{Ing. Amb.}) = 5/14$$

$$P(\text{Agrim.}) = 9/14$$

$$P(X = 2) = \frac{\binom{5}{2} \binom{9}{3}}{\binom{14}{5}}$$

$$P(X = 2) = 0,42$$



Universidad Nacional del Litoral
Facultad de Ingeniería y Ciencias Hídricas

ESTADÍSTICA

Ingeniería Informática

TEORÍA

Mg.Ing. Susana Vanlesberg
Profesor Titular

UNIDAD 4

MODELOS PROBABILÍSTICOS

MODELOS PARA VARIABLES ALEATORIAS CONTINUAS

MODELO EXPONENCIAL

Este modelo surge al considerar el *tiempo hasta* la primera ocurrencia de un evento que pueda ser considerado como proceso de Poisson.

Si la variable aleatoria es ahora el tiempo transcurrido hasta que se verifica la primera ocurrencia, entonces será una variable continua, la probabilidad que T exceda algún valor t es lo mismo que decir que **no** se verificaron ocurrencias en ese intervalo de longitud t, lo que es equivalente a decir que la variable aleatoria **Nº de ocurrencias** de tipo Poisson toma el valor 0:

$$P(T > t) = P(x = 0)$$
$$P(x = 0) = \frac{e^{-\lambda t} \lambda^0}{0!} = e^{-\lambda t}$$

Esto permite obtener la función de distribución de T variable continua:

$$F(t) = P(T \leq t) = 1 - P(T > t) = 1 - e^{-\lambda t} \quad t \geq 0 \quad (1)$$

La función de densidad se obtiene derivando la expresión anterior:

$$f(t) = \lambda e^{-\lambda t} \quad (2)$$

Cumpliendo con las propiedades de estacionariedad e independencia de los procesos de Poisson, $e^{-\lambda t}$ da la probabilidad de que no se verifiquen eventos en algún intervalo de tiempo de longitud t, estando o no el origen en el tiempo 0.

Por ejemplo si el origen estuviese en el momento de la enésima ocurrencia, $e^{-\lambda t}$ representaría la probabilidad de que el tiempo hasta la próxima ocurrencia (n+1) sea mayor que t.

Esto quiere demostrar sencillamente que los tiempos entre ocurrencias de eventos de tipo Poisson son independientes y se distribuyen en forma exponencial.

Características

$$E(T) = \int_0^{\infty} t \lambda e^{-\lambda t} dt$$

$$\text{si } u = t\lambda \quad du = \lambda dt \quad \frac{du}{\lambda} = dt$$

$$\frac{1}{\lambda} \int_0^{\infty} u e^{-u} du = \frac{1}{\lambda} [e^{-u}(-u-1)]_0^{\infty}$$

$$E(T) = \frac{1}{\lambda} \quad (3)$$

Recordar que λ en los procesos de Poisson, representa el número promedio de ocurrencias, aquí, $1/\lambda$ representa el tiempo promedio entre ocurrencias.

$$Var(T) = E(T^2) - E^2(T)$$

$$E(T^2) = \int_0^{\infty} t^2 \lambda e^{-\lambda t} dt$$

$$\text{si } v = t^2 \quad dv = 2t dt$$

$$\lambda e^{-\lambda t} dt = du \quad u = \int_0^{\infty} \lambda e^{-\lambda t} dt = -e^{-\lambda t}$$

$$E(T^2) = [-e^{-\lambda t} t^2]_0^{\infty} - \int_0^{\infty} -e^{-\lambda t} 2t dt =$$

$$= 0 + 2 \int_0^{\infty} e^{-\lambda t} t dt \frac{\lambda}{\lambda} = \frac{2}{\lambda} \left(\frac{1}{\lambda} \right)$$

$$E(T^2) = \frac{2}{\lambda^2}$$

$$Var(T) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda} \right)^2 = \frac{1}{\lambda^2} \quad (4)$$

Una característica de los procesos de Poisson es que no tienen memoria. Esto significa que el comportamiento futuro es independiente de lo registrado en el presente o en el pasado.

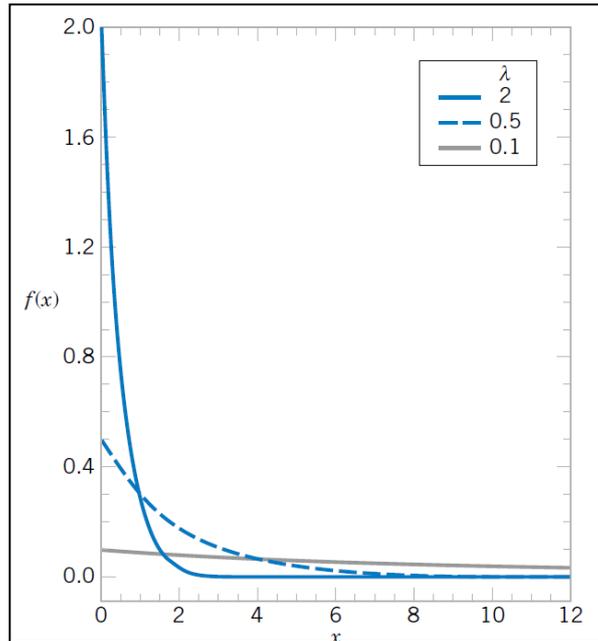


Figura N° 1- Distribución exponencial para distintos valores del parámetro

Ejemplo

El personal de la compañía Onda S.L. usa una Terminal para realizar sus pedidos internacionales. Si el tiempo que cada comercial gasta en una sesión en la Terminal tiene una distribución exponencial de media 36 minutos, encontrar:

- Probabilidad de que un comercial utilice la Terminal 30 minutos o menos.
- Si un comercial ha estado por lo menos 30 minutos en la Terminal, ¿Cuál es la probabilidad de que pase al menos una hora más en la Terminal?

$$1/\lambda = 36 \quad \lambda = 1/36$$

$$a) \quad P(X \leq 30) = \int_0^{30} 1/36 * e^{-1/36*x} dx = 0.565$$

$$b) \quad P((X \geq 60)/(X \geq 30)) = P(X \geq 60) = \int_{60}^{+\infty} 1/36 * e^{-1/36*x} dx = 0.188$$

MODELO NORMAL

Este modelo suele conocerse como **Modelo de las sumas**, ya que la incertidumbre en algunas variables puede ser el resultado de efectos combinados de algunas causas que contribuyen, siendo difícil de separar y observar a cada una. En algunas situaciones si se conoce el mecanismo por el cual las causas individuales afectan a la variable de interés se puede determinar un modelo para la variable resultante sin estudiar en detalle los efectos individuales, particularmente no es necesario conocer la distribución de las causas.

El modelo de la variable resultado, presenta una función de densidad doble exponencial cuya forma es la de una campana. Este hecho está contenido en el Teorema del Límite Central, el cual es uno de los resultados más importantes de la Teoría de Probabilidad. Este teorema será desarrollado más adelante pero se puede hacer referencia a su enunciado: *bajo condiciones generales, cuando el número de variables que intervienen en la suma que origina una variable, es cada vez más grande, la distribución de esta suma tiende a aproximarse al modelo Normal.*

La inmensa importancia práctica del modelo Normal reside entre otras razones en que lo que se plantea en el teorema del Límite Central, puede hacerse sin el conocimiento exacto de:

- las distribuciones marginales de cada variable que interviene en la suma, - de su número, - de su distribución conjunta. Ya que la variación aleatoria en algunos fenómenos naturales se origina de un número de variaciones aditivas, no debe sorprender el hecho que gráficos que pueden aproximarse a este modelo, se observen con frecuencia en la naturaleza.

La función de densidad que caracteriza esta distribución es una doble exponencial con la siguiente forma:

$$f(x) = ke^{-c(x-m)^2}, -\infty \leq x \leq \infty \quad (5)$$

siendo **m** la distancia al centro de la distribución que por la simetría de la distribución es igual a su media μ .

Las constantes k y c se pueden determinar: k por considerar que f(x) para ser función de densidad deberá cumplir la condición de ser igual a uno para todos los valores que la variable puede tomar:

$$\int_{-\infty}^{\infty} k e^{-c(x-\mu)^2} dx = 1$$

$$k \int_{-\infty}^{\infty} e^{-c(x-\mu)^2} dx = 1$$

llamando $a(x - \mu) = y$, $dx = dy$, luego

$$k \int_{-\infty}^{\infty} e^{-c y^2} dy = 1$$

$$\text{siendo } \int_0^{\infty} e^{-a^2 x^2} dx = \frac{\sqrt{\pi}}{2a} \quad a^2 = c \longrightarrow a = \sqrt{c}$$

por simetría se tiene :

$$k 2 \frac{\sqrt{\pi}}{2\sqrt{c}} = 1$$

$$k = \frac{\sqrt{c}}{\sqrt{\pi}}$$

Para completar la función de densidad es necesario evaluar la otra constante, c . Para esto se utiliza la varianza:

$$\text{Var}(x) = \frac{\sqrt{c}}{\sqrt{\pi}} \int_{-\infty}^{\infty} (X - \mu)^2 e^{-c(X-\mu)^2} dx$$

Para resolver esta integral se usará un resultado proveniente de la resolución de integrales impropias dependientes de un parámetro:

$$\int_0^{\infty} e^{-c x^2} dx = \frac{1}{2} \frac{\sqrt{\pi}}{\sqrt{c}} \text{ para } c > 0$$

luego llamando $F(c) = \int_0^{\infty} e^{-c x^2} dx$

y aplicando la siguiente propiedad

$$F(x) = \int_a^{\infty} f(t, x) dt \longrightarrow F'(x) = \int_a^{\infty} \frac{\partial f}{\partial x}(t, x) dt$$

$$F(c) = \int_0^{\infty} f(t, c) dt \longrightarrow F'(c) = \int_0^{\infty} \frac{\partial f}{\partial c}(t, c) dt$$

en este caso origina :

$$F(c) = \int_0^{\infty} e^{-cx^2} dx \quad \text{donde}$$

$$f(t,c) = e^{-ct^2} \quad \text{luego} \quad \frac{\partial f}{\partial c} = -t^2 e^{-ct^2}$$

$$\text{con lo cual } F'(c) = \int_0^{\infty} -t^2 e^{-ct^2} dt$$

$$\text{pero como } F'(c) = \frac{dF}{dc} = \frac{d}{dc} \left(\frac{1}{2} \frac{\sqrt{\pi}}{\sqrt{c}} \right) = -\frac{\sqrt{\pi}}{4} c^{-\frac{3}{2}}$$

resulta :

$$\int_0^{\infty} -t^2 e^{-ct^2} dt = -\frac{\sqrt{\pi}}{4} c^{-\frac{3}{2}}$$

$$\text{luego} \quad \int_{-\infty}^{\infty} t^2 e^{-ct^2} dt = \frac{\sqrt{\pi}}{2} c^{-\frac{3}{2}}$$

$$\text{y} \quad \frac{\sqrt{c}}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^2 e^{-ct^2} dt = \frac{\sqrt{c}}{\sqrt{\pi}} \frac{\sqrt{\pi}}{2} c^{-\frac{3}{2}} = \frac{c^{-1}}{2}$$

$$\text{Var}(X) = \frac{1}{2c} \quad \text{de donde}$$

$$c = \frac{1}{2\sigma^2}$$

Esto permite completar la expresión de la función de densidad:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty \leq x \leq \infty \quad (6)$$

La función acumulativa será a partir de esta función anterior, la siguiente:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt \quad (7)$$

μ y σ son sus parámetros. σ desvío estándar y μ la media. Así es que la función queda completamente definida si se conocen estos dos valores.

- $[(x-\mu)/\sigma]^2$: es la parte práctica de la función ya que contiene un valor determinado de la variable y a los parámetros. El hecho de estar elevado al cuadrado hace que dos valores distintos de la variable, que tengan la misma desviación absoluta de la media μ , van a tener igual valor de la función de densidad. Esto muestra claramente la característica de simétrica alrededor de μ , que posee la distribución Normal. Al ser este exponente negativo, cuanto mayor es la desviación de x respecto de μ , será menor la densidad de probabilidad de x , esto es lo que se observa en las colas de la distribución. Ahora cuando x coincide con μ , el exponente es cero y el valor de la función de densidad es máxima en ese punto e igual a $1/\sigma\pi^{1/2}$, o sea que $x=\mu$ es el valor del modo.

- El valor de μ hace desplazar al modelo hacia la izquierda o la derecha, mientras que el valor de σ cambia su forma sin desplazarlo. Este efecto se observa en las siguientes figuras:

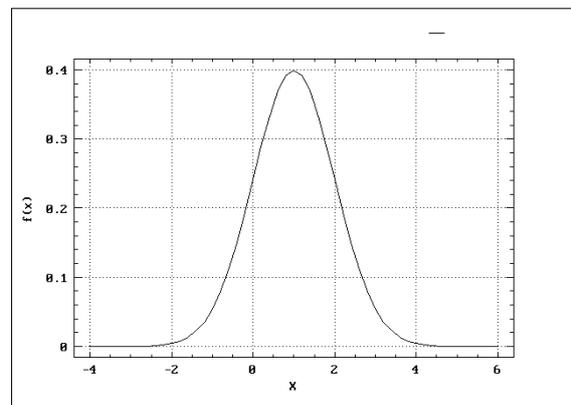
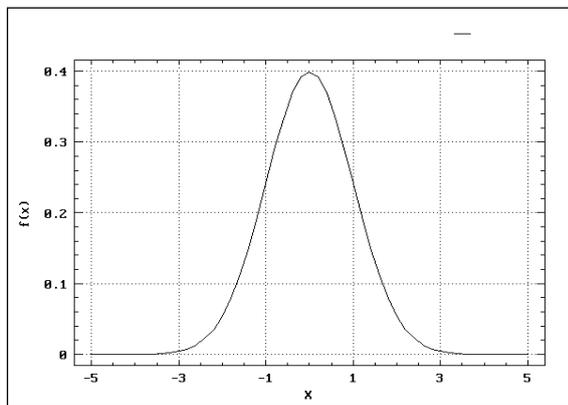


Fig. N° 2 - Dist. Normal con $\mu=0$ y $\sigma=1$ Fig. N° 3 - Dist. Normal con $\mu=1$ y $\sigma=1$

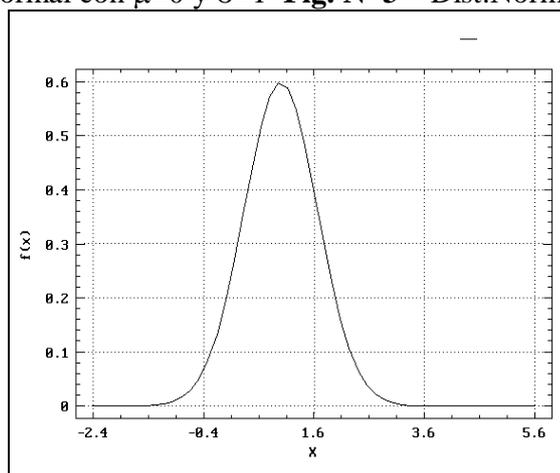


Fig. N° 4 - Distribución Normal con $\mu=1$ y $\sigma=2/3$

- Debido a que la distribución Normal tiene amplitud infinita, su curva nunca toca el eje x, la probabilidad de un intervalo muy alejado de μ será prácticamente despreciable y el 99% de su área está encerrada por $\mu \pm 3\sigma$.

En las siguientes figuras se observa el efecto de tomar distintos valores de μ y σ :

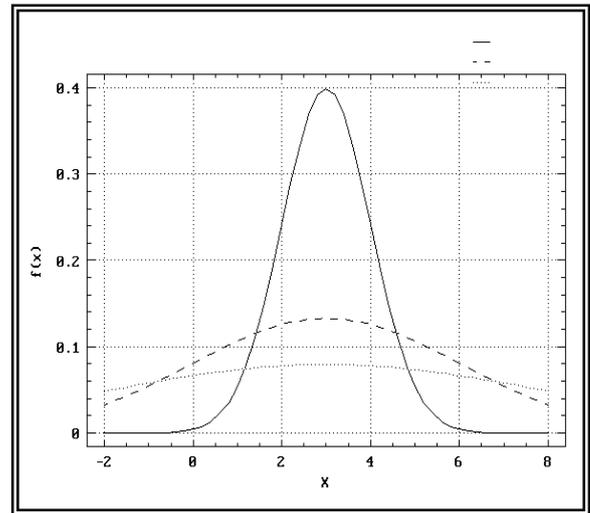
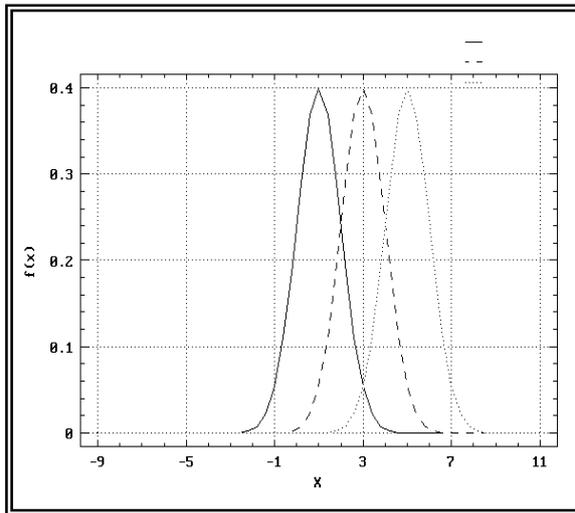


Fig. N° 5 - Curvas Normales: distintos μ e igual σ **Fig. N° 6** - Curvas Normales: igual μ y distintos σ

Características

$$E(X) = \int_{-\infty}^{\infty} xf(x) dx =$$

$$= \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2} dx =$$

siendo $\frac{x-\mu}{\sigma} = y$ $x = \sigma y + \mu$ $dx = \sigma dy$

$$E(X) = \int_{-\infty}^{\infty} (\sigma y + \mu) \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{1}{2} y^2} \sigma dy =$$

Integrando por partes,

$$\begin{aligned}
&= \frac{\sigma}{\sqrt{2\pi} \sigma} \int_{-\infty}^{\infty} (\sigma y + \mu) e^{-\frac{y^2}{2}} dy = \\
&= \frac{1}{\sqrt{2\pi}} \left[\int_{-\infty}^{\infty} (\sigma y) e^{-\frac{y^2}{2}} dy + \int_{-\infty}^{\infty} \mu e^{-\frac{y^2}{2}} dy \right] = \\
&= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-\frac{y^2}{2}} dy + \mu \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \\
&= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} d\left(\frac{y}{2}\right) + \mu = \\
&= \frac{\sigma}{\sqrt{2\pi}} \left[-e^{-\frac{y^2}{2}} \right]_{-\infty}^{\infty} + \mu = 0 + \mu \\
&E(X) = \mu \tag{8}
\end{aligned}$$

$$\begin{aligned}
\text{Var}(X) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\
\frac{x-\mu}{\sigma} &= y \quad dy = \frac{dx}{\sigma} \\
\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma^2 y^2 e^{-\frac{1}{2}y^2} \sigma dy &= \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-\frac{1}{2}y^2} d\left(\frac{y^2}{2}\right) \\
u = y \quad dv &= e^{-\frac{1}{2}y^2} d\left(\frac{y^2}{2}\right) \Rightarrow v = -e^{-\frac{y^2}{2}} \\
\frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy &= \sigma^2 \left\{ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2} dy \right\} \\
\text{Var}(X) &= \sigma^2 \tag{9}
\end{aligned}$$

Otras características del modelo normal que merecen destacarse:

-Muchas variables aleatorias continuas se distribuyen en forma aproximadamente normal. Es de destacar que los errores de mediciones repetidas de alguna dimensión particular, se dice que tienen distribución normal. Esto se debe a que cualquier medida se supone que es igual a un valor verdadero más un error; este error puede considerarse como el resultado de un gran conjunto de factores que están presentes en ese momento, y cada factor ejerce un pequeño efecto en la magnitud y sentido del error. Los errores actúan independientemente y con igual fuerza para aumentar o disminuir el valor de la medición observada, y a largo plazo se anulan. Esto hace considerar a los errores de medida como distribuidos normalmente, con valor medio cero, y se los suele denominar "errores al azar".

-Este modelo sirve como buena aproximación de modelos discretos, como el Binomial o el de Poisson, bajo circunstancias especiales.

- El supuesto de normalidad de las poblaciones permite obtener buenos resultados de métodos elaborados bajo este supuesto, aunque en realidad no se cumpla en forma estricta.

-Muchos estadísticos calculados a partir de grandes muestras se distribuyen en forma aproximadamente normal, lo cual facilita el trabajo de inferencia estadística.

Modelo Normal estándar

La distribución normal se encuentra tabulada y existen, además, rutinas computacionales que permiten trabajar con ella.

Para fines prácticos y para ganar en eficiencia y rapidez, el concepto de distribución estándar es muy importante.

Se denomina modelo estándar a aquel cuya media es 0 y su desvío es uno. La variable $z=(x-\mu)/\sigma$ se denomina *variable aleatoria estandarizada*, cuya media es 0 y su desvío es unitario.

La función de densidad del modelo estandarizado es:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad -\infty < z < +\infty \quad (10)$$

y su función acumulativa es la siguiente:

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{1}{2}t^2} dt \quad (11)$$

Características de este modelo:

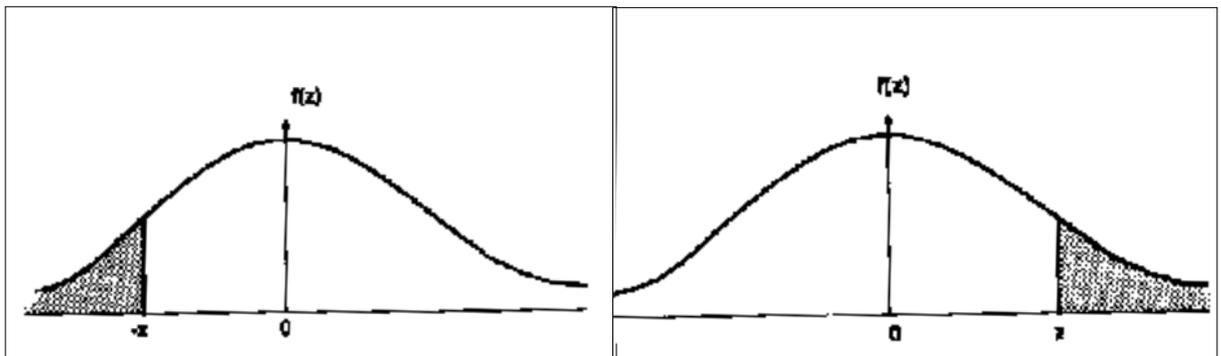
$$E(z) = E\left(\frac{x - \mu}{\sigma}\right) = \frac{1}{\sigma} E(x - \mu) = \frac{1}{\sigma} [E(x) - E(\mu)] = \frac{1}{\sigma} (\mu - \mu) \quad (12)$$

$$Var(z) = Var\left(\frac{x - \mu}{\sigma}\right) = \frac{1}{\sigma^2} Var(x - \mu) = \frac{1}{\sigma^2} Var(x) - 0 = \frac{\sigma^2}{\sigma^2}$$

$$Var(\mu) = E[\mu - E(\mu)]^2 = E(\mu - \mu)^2 = 0$$

$$Var(z) = 1 \quad (13)$$

Para el manejo de las tablas, a veces es necesario tener en cuenta el carácter simétrico de la distribución, ya que algunas están elaboradas solamente para valores positivos. Para obtener probabilidades o densidades de valores negativos de la variable se procede de la siguiente manera:



$$P(Z \leq -z) = P(Z \geq z) = 1 - P(Z \leq z)$$

Esta simetría del modelo normal respecto a su media implica que los momentos centrados de orden impar sean cero. Los de orden par pueden considerarse a partir de la siguiente expresión:

$$\mu_n = E(x - \mu)^n = \frac{n!}{2^{\frac{n}{2}} \left(\frac{n}{2}\right)!} \sigma^n \quad n = 2, 4, \dots \quad (14)$$

La asimetría será, por lo tanto, igual a cero:

$$As = \frac{\mu_3}{\sigma^3} = 0 \quad (15)$$

y la kurtosis, en función de la expresión general de momentos centrados de orden par, será:

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

$$\mu_4 = \frac{4!}{2^2 \left(\frac{4}{2}\right)!} \sigma^4 = 3\sigma^4$$

$$K = 3 \quad (16)$$

Es por esta característica que suelen compararse los coeficientes de kurtosis de distintos modelos con el de la normal, permitiendo clasificarlos según sean mayores, menores o iguales que 3.

Ejemplo

Una máquina fabrica tornillos cuyas longitudes se distribuyen normalmente con media 20 mm y varianza 0.25 mm. Un tornillo se considera defectuoso si su longitud difiere de la media más de 1 mm. Los tornillos se fabrican de forma independiente. ¿Cuál es la probabilidad de fabricar un tornillo defectuoso?

$$\sigma = 0.5$$

$$\sigma^2 = 0.25$$

$$x \longrightarrow N(20, 0.5)$$

$$P(\text{def.}) = P(x > 21) + P(x < 19) = P\left(z > \frac{21 - 20}{0.5}\right) + P\left(z < \frac{19 - 20}{0.5}\right) = P(z > 2) + P(z < -2) =$$

$$[1 - P(z \leq 2)] + [1 - P(z \leq 2)] = [1 - 0.9773] + [1 - 0.9773] = 0.0452$$

Distribución de la suma de variables aleatorias Normales

El Teorema del Límite Central (que se expondrá más detalladamente más adelante), expresa que la suma de un número dado de variables aleatorias tiende a distribuirse en forma Normal. Es de esperar que si se suma un número de variables aleatorias independientes y Normales cada una, la suma será distribuida en forma Normal.

$$x_1 \sim N(\mu_1, \sigma_1)$$

$$x_2 \sim N(\mu_2, \sigma_2)$$

.

$$x_n \sim N(\mu_n, \sigma_n)$$

$$U = \sum_{i=1}^n x_i$$

con parámetros

$$E(U) = E\left(\sum_{i=1}^n x_i\right) = \sum_{i=1}^n E(x_i) = \sum_{i=1}^n \mu_i$$

$$\sigma^2(U) = \sigma^2\left(\sum_{i=1}^n x_i\right) = \sum_{i=1}^n \sigma^2(x_i) = \sum_{i=1}^n \sigma_i^2$$

$$\text{entonces } U \sim N\left(\sum \mu_i; \sqrt{\sum_{i=1}^n \sigma_i^2}\right)$$

Aproximación del modelo Binomial al Normal

En el esquema que corresponde a una variable aleatoria con distribución Binomial, cuando n aumenta y p no varía, la variable aleatoria

$$\frac{x - np}{\sqrt{npq}}$$

se distribuye aproximadamente Normal:

$$P(X \leq x_0) = F_N\left(\frac{x_0 - np}{\sqrt{npq}}\right)$$

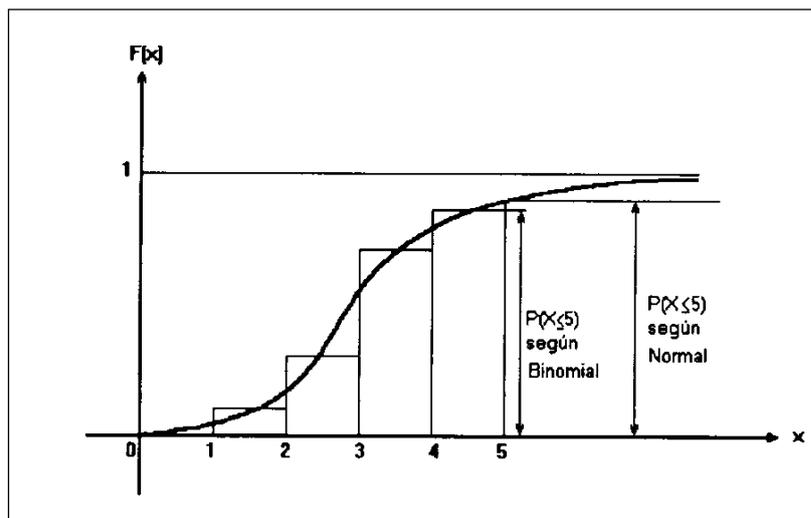
siendo x_0 , un valor cualquiera de x .

El primer miembro de la última ecuación pertenece a la distribución Binomial, cuya función de distribución es escalonada. El segundo miembro pertenece a la distribución Normal que es continua. A pesar de esta diferencia, para grandes valores de n ambas pueden considerarse iguales para cada valor, ya que la función continua pasa por la mitad de cada escalón de la función discreta. Esto es una buena aproximación para valores de

$n \cdot p > 5$, cuando $p \leq 1/2$ y $n \cdot q > 5$ con $p > 1/2$. Estas aproximaciones son válidas si n es pequeño introduciendo un factor de corrección por continuidad cuyo valor es igual a $1/2$.

En la figura siguiente se observa que una mejor aproximación a $P(X \leq 5)$ se obtiene por tomar la ordenada de la distribución continua $1/2$ unidad a la derecha de 5. En general:

$$P(X \leq x_0) = F_N \left(\frac{x_0 + \frac{1}{2} - np}{\sqrt{npq}} \right)$$



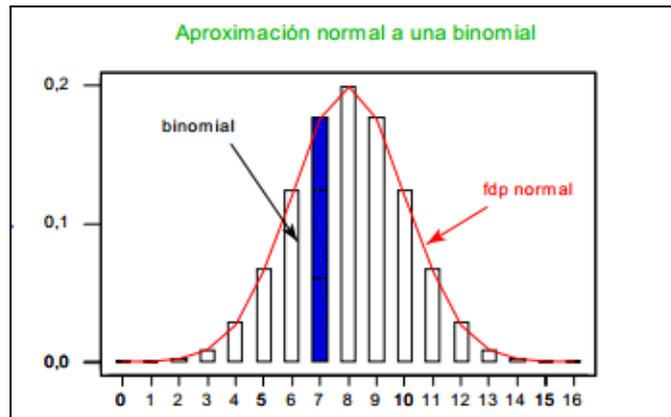


Figura N° 7 – Aproximación Binomial con la Normal

DISTRIBUCIONES RELACIONADAS CON LA DISTRIBUCIÓN NORMAL

Distribución Chi-cuadrado

Este modelo describe la distribución de la suma de los cuadrados de v variables aleatorias independientes, con distribución $N(0,1)$:

$$V = \sum_{i=1}^v X_i^2 \sim \chi_v^2 \quad (17)$$

Si las variables que intervienen en la suma no fuesen $N(0,1)$, se las debería estandarizar, con lo cual se obtendría:

$$V = \sum_{i=1}^v \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2 \sim \chi_v^2 \quad (18)$$

V es también una variable aleatoria, ya que es el resultado de la suma de variables aleatorias. Por ser una suma de cuadrados, varía desde 0 a infinito.

La función de densidad surge de considerar la función $N(0,1)$ de cada componente al cuadrado, y es la siguiente:

$$f(\chi^2) = \begin{cases} \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} \chi^{2\left(\frac{\nu-2}{2}\right)} e^{-\frac{\chi^2}{2}}, & \text{para } \chi^2 > 0 \\ 0, & \text{en otro caso} \end{cases} \quad (19)$$

siendo Γ la función gamma y ν el número de variables al cuadrado que intervienen en la suma, y se denominan "grados de libertad".

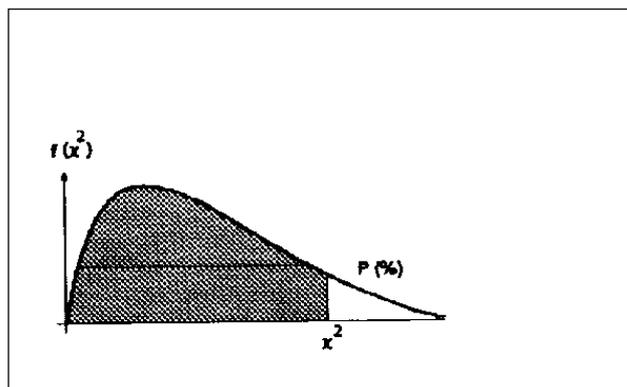


Figura N° 8 – Distribución chi- cuadrado

Las características de esta distribución son:

$$\begin{aligned} E(\xi^2) &= \nu \\ \sigma^2(\xi^2) &= 2\nu \end{aligned} \quad (20)$$

Las tablas permiten evaluar probabilidades del siguiente tipo:



Propiedades de la distribución chi-cuadrado

- 1.- La variable solo puede tomar valores positivos.
- 2.- Es asimétrica.
- 3.- Depende del parámetro ν (grados de libertad).
- 4.- Su esperanza matemática es ν , y su varianza, 2ν .
- 5.- *Propiedad aditiva o reproductiva* : Si χ^{2n} y χ^{2m} son dos variables Chi cuadrado con n y m grados de libertad respectivamente, independientes entre sí, entonces la suma de las dos variables es una variable Chi-cuadrado con $n+m$ grados de libertad. Esto se puede generalizar a la suma de cualquier número de variables Chi-cuadrado, independientes.
- 6.- Al aumentar el número de grados de libertad, la distribución Chicuadrado se aproxima asintóticamente a una distribución normal.

Distribución t de Student

Este modelo surge como el cociente entre una variable $N(0,1)$ y la raíz cuadrada de una variable aleatoria distribuida χ^2 dividida por sus grados de libertad, siendo estas dos variables independientes:

$$t_\nu = \frac{x}{\sqrt{\frac{V}{\nu}}}, \quad \begin{cases} x \sim N(0,1) \\ V \sim \chi_\nu^2 \end{cases} \quad (21)$$

El rango de la variable t varía entre $-\infty$ y $+\infty$, y su función de densidad surge de las distribuciones de las variables aleatorias componentes:

$$f(t) = \frac{1}{\sqrt{\pi\nu}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad \text{para } -\infty < t < \infty \quad (22)$$

siendo ν los grados de libertad asociados a la componente χ^2

Sus características son:

$$\begin{aligned} E(t) &= 0 \\ \sigma^2(t) &= \frac{\nu}{\nu-2}, \quad \nu > 2 \end{aligned} \quad (23)$$

La gráfica es parecida a la de la distribución Normal, y es simétrica respecto del cero. Para grandes valores de ν se la puede aproximar a la Normal.

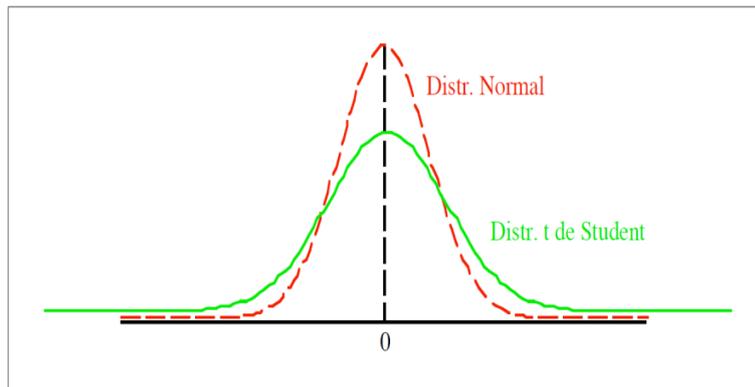


Figura N° 9 - Distribución de Student

Existen tablas que brindan probabilidades para distintos valores de la variable y algoritmos computacionales que permiten obtener probabilidades con igual o mayor grado de precisión que las tablas.

Propiedades de la distribución "t"

- 1.- Depende de un único parámetro, el número de grados de libertad.
- 2.- El rango de la variable es todo el eje real $(-\infty, +\infty)$.
- 3.- Su gráfica es simétrica respecto al eje de ordenadas OY.
- 4.- El valor $x = 0$ es la media, mediana y moda de la distribución.
- 5.- Al aumentar n , se va haciendo cada vez más apuntada la gráfica de su función de densidad, siendo el límite para $n \rightarrow \infty$ la curva normal tipificada.

Distribución F de Snedecor

Si X e Y son dos variables aleatorias independientes cada una con distribución χ^2 , con ν_1 y ν_2 grados de libertad; luego la distribución F está definida por el siguiente cociente:

$$F_{\nu_1, \nu_2} = \frac{\frac{X}{\nu_1}}{\frac{Y}{\nu_2}} = \frac{\frac{\chi_1^2}{\nu_1}}{\frac{\chi_2^2}{\nu_2}} \quad (24)$$

Como es el cociente entre dos variables aleatorias χ^2 , que son positivas, también será positiva, y su gráfica similar a la de χ^2

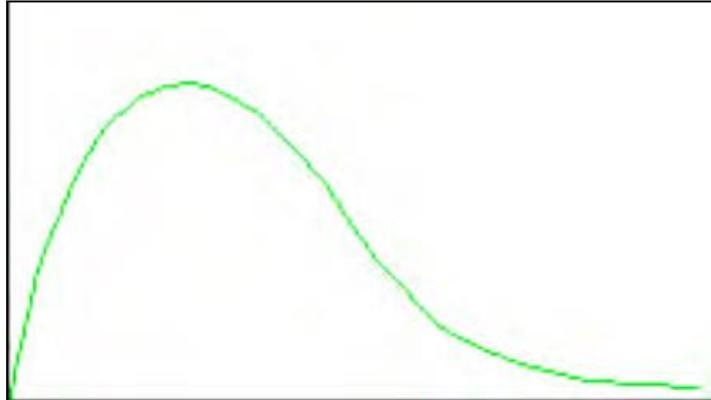


Figura N° 10 - Distribución de Snedecor

Su función de densidad surge del conocimiento de las variables aleatorias ξ^2 que la integran, y es la siguiente:

$$f(F) = \frac{\Gamma\left(\frac{\nu_1 + \nu_2}{2}\right) \nu_1^{\frac{\nu_1}{2}} \nu_2^{\frac{\nu_2}{2}} F^{\frac{\nu_1}{2}-1}}{\Gamma\left(\frac{\nu_1}{2}\right) \Gamma\left(\frac{\nu_2}{2}\right) (\nu_2 + \nu_1 F)^{\frac{\nu_1 + \nu_2}{2}}}, \text{ para } F > 0 \quad (25)$$

Se utiliza fundamentalmente en la parte de inferencia estadística. Existen tablas que brindan valores de probabilidad del siguiente tipo: $P(F > F_p, \nu_1, \nu_2) = P(\%)$. Debido a que la función depende de los grados de libertad, se necesita una tabla a triple entrada para obtener valores tabulados de F que corresponden a distintas probabilidades y a distintos valores de los grados de libertad.

A veces es necesario tener en cuenta la siguiente relación:

$$F_{1-\alpha, \nu_2, \nu_1} = \frac{1}{F_{\alpha, \nu_1, \nu_2}} \quad (26)$$

MODELOS DE VALORES EXTREMOS

En muchas situaciones prácticas especialmente en ingeniería, es de interés trabajar con el mayor o el menor valor de un número de variables aleatorias.

Si la variable y es considerada el máximo de una serie de n variables aleatorias $x_1 \dots x_n$ es posible obtener series y_t que estarán compuestas por los máximos. Es posible así obtener una expresión para la probabilidad de que el máximo sea menor o igual que un valor dado:

$$P(Y \leq y) = F(y) = P(\text{para todo } n \text{ de } x_i \leq y)$$

Si los valores de x son independientes, entonces:

$$\begin{aligned} F(y) &= P(X_1 \leq y) P(X_2 \leq y) \dots P(X_n \leq y) = \\ &= F_{x_1}(y) F_{x_2}(y) \dots F_{x_n}(y) \end{aligned}$$

Si los x_i son idénticamente distribuidos, $F(x)$, luego:

$$F(y) = [F_x(y)]^n$$

y, cuando $n \rightarrow \infty$: se deriva el modelo buscado.

De acuerdo a las características de la distribución inicial se originan tres tipos de modelos asintóticos de valores extremos:

Modelo Tipo I:

Surge de aquellas distribuciones iniciales que no tienen límite superior. El extremo de la curva correspondiente a la función de densidad debe decrecer tan rápidamente como una función exponencial; entonces, valores extremos provenientes de una distribución normal, log-normal, gamma, pueden ser ajustados por un Modelo Tipo I de máximos. En cambio, si la distribución inicial no es limitada en la dirección de los mínimos, se origina un Modelo Tipo I de mínimos.

Modelo Tipo II:

Este modelo surge de aquellas distribuciones iniciales ilimitadas, y que poseen un número finito de momentos.

Modelo Tipo III:

Este modelo surge cuando la distribución inicial está limitada en la dirección del valor extremo, es así que la distribución de valores mínimos provenientes de distribuciones log-normal, gamma y beta pueden ajustarse por un Modelo Tipo III.

Existen una serie de condiciones que se deben establecer en el desarrollo del modelo:

- Las observaciones de las que se extraen los valores extremos deben ser independientes.
- Las observaciones deben ser hechas bajo las mismas condiciones, es decir, las distribuciones iniciales y los parámetros que contienen deben ser los mismos.
- El número de observaciones n de las cuales se extraen los valores extremos debe ser grande; en algunas situaciones, día y años son unidades naturales de periodicidad.

Modelo Tipo III -Weibull

Este modelo se origina con las mismas consideraciones hechas para el anterior, a las cuales se les agrega las siguientes:

- La distribución de los X_i está limitada superiormente por un valor ω .
- La distribución de X_i es del siguiente tipo:

$$P(X \leq x) = 1 - c(\omega - x)^k$$

con $x \leq \omega, k > 0$

De esto se obtienen las funciones para el mayor valor de los X_i :

$$F(x) = e^{-\left(\frac{\omega-x}{\omega-\mu_0}\right)^k}, \quad x < \omega$$
$$f(x) = \frac{k}{\omega - \mu_0} \left(\frac{\omega-x}{\omega-\mu_0}\right)^{k-1} e^{-\left(\frac{\omega-x}{\omega-\mu_0}\right)^k} \quad (37)$$

Esta forma del modelo no es muy aplicada en muchas ramas de las ciencias, quizá por el límite superior que presenta. Una transformación de la variable resulta en el conocido modelo de Weibull. Si se trabaja con el valor negativo de la variable $z = -x$, estando x definida para valores menores que ω , z estará definida para valores mayores que $-\omega$. Llamando $-\omega = \varepsilon$, y observando que las probabilidades de no excedencias serán ahora de excedencias, la función de probabilidad estará dada por:

$$P(Z \geq z) = e^{-\left(\frac{z-\varepsilon}{\mu_0-\varepsilon}\right)^k}$$

Por lo tanto, la función de distribución será:

$$P(Z \leq z) = 1 - e^{-\left(\frac{z-\varepsilon}{\mu_0-\varepsilon}\right)^k} \quad (38)$$

$$f(z) = \frac{k}{\mu_0-\varepsilon} \left(\frac{z-\varepsilon}{\mu_0-\varepsilon}\right)^{k-1} e^{-\left(\frac{z-\varepsilon}{\mu_0-\varepsilon}\right)^k}$$

Al trabajar con los valores negativos de una serie y maximizarlos, se estará obteniendo como extremo un valor mínimo, ya que el mayor valor de una serie negativa es el menor valor en valor absoluto.

Los parámetros de este modelo son μ_0 , k y ε , siendo este último el límite inferior de los valores mínimos.

Las características son las siguientes:

$$E(x) = \mu = \varepsilon + (\mu_0 - \varepsilon) \Gamma\left(1 + \frac{1}{k}\right) \quad (39)$$

$$\text{Var}(x) = \sigma^2 = (\mu_0 - \varepsilon)^2 \left[\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right) \right] \quad (40)$$

Generalmente es posible adoptar $\varepsilon = 0$, lo cual simplificaría las expresiones del modelo:

$$F(z) = 1 - e^{-\left(\frac{z}{\mu_0}\right)^k} \quad (41)$$

$$f(z) = \frac{k}{\mu_0} \left(\frac{z}{\mu_0}\right)^{k-1} e^{-\left(\frac{z}{\mu_0}\right)^k}$$

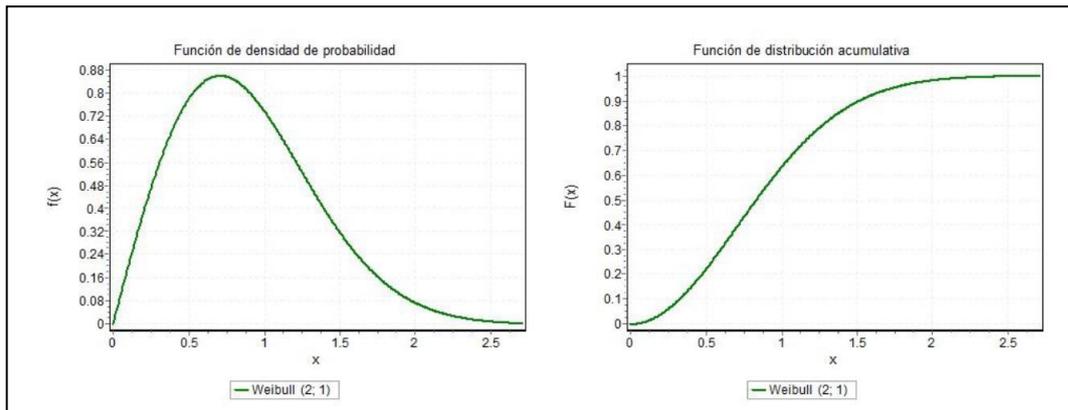


Figura N° 16 – Funciones de la distribución de Weibull para distintos valores de sus parámetros

La primera distribución representada es la que tiene como parámetro de forma igual a 2 y de escala igual a 1; al no haber parámetro de localización, este se supone igual a 0, con lo cual la densidad y la distribución existen para valores mayores que ese valor. A la vista de la gráfica de la función de densidad, se puede deducir que en este caso la distribución es asimétrica positiva, mientras que de la representación de la función de distribución se puede deducir que el crecimiento suele ser constante hasta que $x= 1.5$, a partir de donde empieza a decaer ligeramente para crecer cada vez menos.

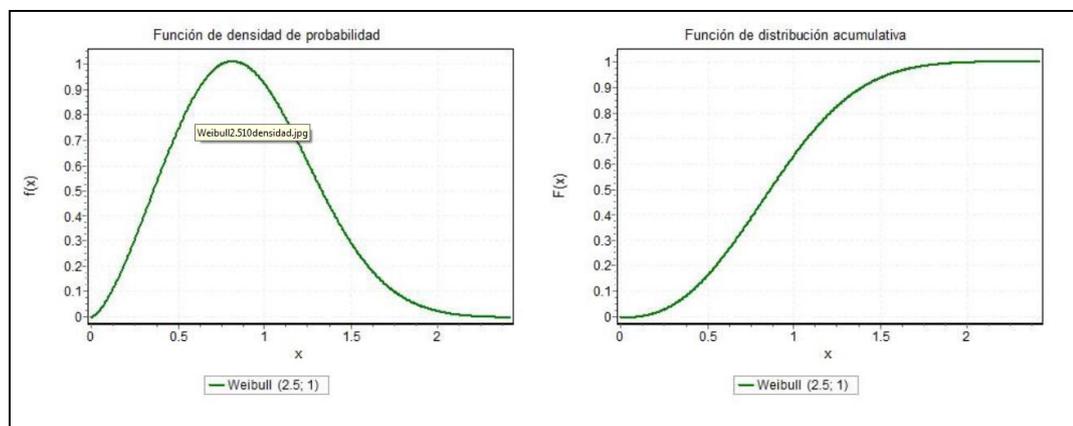


Figura N° 17 – Funciones de la distribución de Weibull para distintos valores de sus parámetros

En este caso, el parámetro de forma vale 2.5, mientras que el de escala es igual a 1. La distribución es asimétrica positiva al estar la cola a la derecha y los valores con mayor probabilidad más a la izquierda.

Se observa que se tienen gráficas de forma flexible a medida que se cambian los parámetros con lo cual se adaptan a una buena cantidad de fenómenos tanto el modelo tipo I como el tipo III.

