UNIVERSIDAD TECNOLÓGICA NACIONAL

FACULTAD REGIONAL DE INGENIERÍA “DELTA”
(Campana)

Probabilidad y Estadística

(Química, Electricidad, Sistemas, Mecánica)

2023

Introducción

La palabra “probabilidad” remite a la idea intuitiva de prueba. La característica singular de una prueba es que no se tiene un conocimiento a priori acerca del resultado, es decir, previamente a la ejecución de la prueba. Por lo tanto, el problema que se presenta es el de la incertidumbre en relación con el resultado y sus consecuencias.

El problema de la incertidumbre es uno de los más frecuentes y realistas. Se relaciona de manera directa con la idea de azar y, en consecuencia, con la experiencia de los juegos de azar. Históricamente se comenta que fueron estos juegos los que ofrecieron el primer marco teórico en el que se desarrollaron por primera vez, en el siglo XVII, los primeros cálculos de probabilidad en manos de Blas Pascal.

La estadística tiene un origen etimológicamente distinto, relativo a los asuntos de estado. El problema se plantea en términos del procesamiento de grandes conjuntos de información que el estadista requiere para tomar decisiones.

Se opone la idea de probabilidad a la noción de determinismo en el sentido de un proceso que tiene un resultado absolutamente conocido. En tal situación no hay dificultad alguna en definir el curso de acción a seguir. Estrictamente no se está tomando una decisión sino una “determinación”, debido a que el resultado de la acción está claramente preestablecido. Es justamente en el ámbito de las probabilidades en el cual se trata el problema de la “decisión”. En toda decisión se presentan opciones y un riesgo de error de no haber tomado la decisión correcta. El problema se resume así a evaluar el nivel de riesgo y a presentar y procesar conjuntos de información que contribuyan a facilitar la toma de decisión. En tal sentido el cálculo de probabilidades recurre a la matemática como auxiliar para la evaluación de riesgo y toma de decisión.

Si nos ubicamos en el contexto histórico, puede verse que, durante el siglo XV, en Europa occidental, los problemas de Dios comenzaron a ser desplazados por los problemas del hombre. En tal sentido, la determinación de los designios divinos comienza a ser reemplazada por la incertidumbre ante las acciones humanas. Durante el siglo XVI nuevos problemas acometen sobre la civilización occidental, nuevos mundos, nuevas rutas. Una visión diferente del cielo a través de un ojo de cristal, leyes escritas sobre planos inclinados con el tiempo marcado por el reloj del corazón. Pero progresivamente la matemática se tornó una herramienta para la comprensión del mundo a través de la razón. Entre las razones, la “razón de estado” se impone como criterio para la toma de decisión entre los estadistas, que requerían información para tomar sus decisiones. En este ámbito se desarrolla el cálculo de probabilidades que mucho después se ofrecerá como sustento teórico para la estadística. el desarrollo del cálculo de probabilidades, durante el siglo XVII, coincide con el de varios otros campos de las matemáticas, las ciencias y los problemas de Estado.

Pero el punto de partida fueron los juegos de azar como la forma más pura de incertidumbre, a la vez la más sencilla para iniciar el desarrollo del cálculo de probabilidades. A este enfoque del problema se lo suele interpretar como “empírico” en el sentido de partir de la interpretación de la prueba, del experimento, la evaluación de los posibles resultados y la mayor o menor factibilidad de que algunos de los resultados se obtengan al cabo de la realización de la experiencia. Ante este enfoque más práctico del cálculo de probabilidades se plantea una visión más formal, matemática y finalmente axiomática del problema, en realidad desarrollada durante el siglo XVIII.

Unidad Nº 1

Generalidades sobre Combinatoria y Cálculo de probabilidades

El problema del conteo

Una primera noción empírica que apunta hacia una definición matemática de la probabilidad parte de la idea de "experimento aleatorio", entendiendo por tal la descripción clara, completa, detallada, de los procedimientos a realizar con el fin de obtener un resultado que no se conoce a priori, pero realizable en idéntica manera por otro experimentador, de manera tal que el resultado no dependa del sujeto que realiza la experiencia sino de las características de la experiencia misma y del azar. Precisamente el azar es lo que introduce la aleatoriedad en el problema que se plantea. Por ejemplo, tirar una moneda, medir la altura de una puerta o describir numéricamente el estado de todas las variables de un avión en vuelo. Esto nos ofrece tres ejemplos claramente diferentes. En primer lugar, un experimento con dos resultados claramente diferenciados (cara y seca), pero sólo dos si se excluye la posibilidad de la rotura de la moneda y otras alternativas, es decir un experimento con un conjunto discreto de resultados posibles. El segundo caso hace referencia a una medición (*), con lo cual se plantea toda medida es el resultado de un experimento aleatorio y, por lo tanto, susceptible de ser tratado en términos probabilísticos con un rango continuo e infinito de resultados posibles. Pero una única medición al fin. El tercer ejemplo nos plantea un conjunto de variables que deben ser medidas en forma simultánea, algunas discretas, otras continuas, pero todas describen conjuntamente el estado del sistema.

Es claro que el tercer ejemplo es el más complejo, el segundo nos plantea el problema del continuo y del infinito, y sólo el primer ejemplo se ofrece, de una manera simple, como un conjunto de resultados posibles claramente diferenciados. Este experimento (tirar una moneda) es totalmente conocido de antemano, si se excluyen caídas laterales, roturas, pérdidas, es decir, si se idealizan las dos alternativas obvias como las únicas posibles, se supone que la moneda está perfectamente equilibrada y que se la ha dejado caer sin ninguna intencionalidad y sometiéndose al puro azar. Lo único que no es conocido entonces es el resultado final del experimento, es decir, si se obtendrá Cara o Ceca.

Con esas consideraciones acerca de la moneda, la realización del experimento y conocimiento de las reglas del juego, puede evaluarse el número de resultados posibles de tirar una monead como "dos" (cp = 2 o número de casos posibles es igual a dos) y también el de resultados favorables si se ha apostado a uno de los dos, por ejemplo, a Cara. En tal caso hay sólo un resultado favorable (cf = 1 o número de casos favorables es igual a uno). Es natural expresar entonces la idea intuitiva o empírica de probabilidad como "uno de dos", "la mitad", o 0.5, a veces expresado en porcentaje como 50%. Es decir que la noción intuitiva nos conduce a "definir" la probabilidad como un cociente entre los casos favorables y los posibles. Más adelante se desarrollará este punto y analizará sus consecuencias. Por el momento podemos abordar un primer problema, consecuencia del anterior, y primera pregunta que puede hacerse a la matemática para contribuir a tratar el problema de la incertidumbre. Este problema consiste en facilitar el conteo.

Es claro que los resultados posibles en una moneda son dos, en un dado son seis, pero ¿en dos dados tirados simultáneamente?, ¿en un complejo juego de cartas? Ante esta dificultad la matemática responde con un conjunto de técnicas que suelen conocerse como "técnicas de conteo".

(*) En una medición se distinguen en primera instancia cuatro elementos, un objeto a medir, un instrumento de medida, un patrón de referencia que define una unidad, y un procedimiento de medición que describe cómo interactúan los elementos mencionados para dar como resultado un valor medido. El procedimiento de medición es lo que en teoría de probabilidades se llama "experimento aleatorio".

Noción intuitiva de probabilidad

Introduciéndonos en el cálculo de probabilidades y a la noción intuitiva de probabilidad con más detalle, analizamos primero los elementos mínimos de un juego de azar. Para ello usaremos como referencia dos ejemplos muy simples, pero suficientemente claros para comprender los conceptos básicos. Sean éstos una y dos monedas, y un dado.

Al tirar una moneda decimos que tenemos dos opciones posibles como resultados válidos, que apostaremos a una como favorable, y que la posibilidad de éxito es del cincuenta por ciento. Lo que hemos hecho es simplemente considerar las reglas del juego válidas para una moneda equilibrada en la que la cara y la ceca son distinguibles.

Que sea equilibrada –que no esté doblada o con el peso desbalanceado– asegura que las dos caras de la moneda tengan la misma posibilidad de salir. Que sean distinguibles quiere decir que la “cara” del rey y la “ceca” donde la moneda fue acuñada tengan un dibujo diferente y no haya confusión acerca del resultado.

El “experimento aleatorio” consiste en tirar la moneda, pero no de una manera deliberada para que salga el resultado deseado, sino forzando la aleatoriedad o el azar en el resultado final. Para ello se la tira al aire y deja que gire y golpee varias veces antes de detenerse. En el juego de azar forzamos así la aleatoriedad de modo que no pueda manipularse el resultado y haya incertidumbre total. De esta manera podemos decir que ambos resultados son igualmente posibles y, si hemos considerado uno como favorable, la idea intuitiva que suele expresarse como “una de dos posibilidades a favor”, se traduce naturalmente en un formato matemático como

$$probabilidad\ de\ \text{é}xito = \frac{n\text{ú}mero\ de\ casos\ favorables}{n\text{ú}mero\ de\ casos\ posibles}$$

En un planteo más simbólico, si llamamos p al número “probabilidad de éxito” y abreviamos las referencias a casos favorables y posibles, escribimos

$$p = \frac{cf}{cp}$$

En el ejemplo de tirar un dado, se asume que debe estar perfectamente equilibrado, es decir que debe ser un cubo perfecto y el centro de gravedad coincida con el centro geométrico, las caras deben ser distinguibles, y el experimento debe realizarse de modo que garantice la aleatoriedad.

El experimento correspondiente a tirar dos monedas es el más sencillo dentro de un nivel superior en complejidad. Notemos que hay cuatro resultados diferentes de acuerdo con las alternativas

Cara	Cara
Cara	Ceca
Ceca	Cara
Ceca	Ceca

Vemos que en este caso el ordenamiento en la secuencia de resultados conduce a cuatro opciones posibles a partir de un experimento realizado con dos monedas. Es fácil ver que, si el experimento se realiza con tres monedas, habrá ocho resultados posibles.

En relación con los resultados favorables, si se considera que “cara” es un resultado favorable entre los dos posibles de tirar una moneda, vemos que, a partir de la definición informal anterior,

$$p = \frac{cf}{cp} = \frac{1}{2} = 0{,}5 = 50\%$$

El cálculo nos lleva inmediatamente a expresar la probabilidad como una fracción, como un número decimal, o en porcentaje.

Al tirar un dado, la probabilidad de que quede en posición superior la cara que contiene dos puntos, el número “2”, se obtiene dividiendo un caso favorable (el 2) entre seis caras o casos posibles, de modo que

$$p = \frac{cf}{cp} = \frac{1}{6} = 0{,}1666\ldots \cong 0{,}167 = 16{,}7\%$$

Podemos preguntarnos cuánto vale la probabilidad de obtener dos caras al tirar una moneda dos veces, y vemos a partir de la tabla que vale

$$p = \frac{cf}{cp} = \frac{1}{4} = 0{,}25 = 25\%$$

En cambio, si preguntamos “cuánto vale la probabilidad de obtener una cara” al tirar dos monedas, vemos que hay dos resultados favorables entre tres posibles. Si llamamos éxito a obtener al menos una cara, habrá tres resultados posibles.

Si al tirar un dado nos preguntamos por la probabilidad de obtener un número par vemos que hay tres resultados favorables. Si pedimos que se obtenga un número menor que tres habrá dos resultados favorables. Si debe ser el número par y menor que tres, habrá sólo un resultado favorable: el “2”. Pero si queremos que sea par o menor que tres, serán favorables el “1”, el “2”, el “4” y el “6”.

Queda como ejercicio analizar estas afirmaciones y calcular las probabilidades en formato de fracción, de número decimal y de porcentaje.

Por otra parte, vemos que no puede haber casos favorables negativos. Eventualmente puede no haber ningún caso favorable ($cf = 0$), con lo cual la probabilidad de éxito será nula. Tomemos como ejemplo pedir que salgan tres caras al tirar dos monedas. O bien podemos plantear una situación en la que todos los casos son favorables; sea por ejemplo pedir menos de tres caras al tirar dos monedas. Es claro que habrá cuatro casos favorables, tantos como posibles. En el primer ejemplo, la probabilidad de éxito es nula, de modo que se tiene la seguridad absoluta de fracaso. El evento es entonces “determinístico” y no “probabilístico”. En el segundo ejemplo, la probabilidad de éxito vale “uno”, o “cien por ciento”. Se trata de otro evento determinístico con seguridad absoluta de éxito.

En ambos casos ($p = 0$) y ($p = 1$) se tiene seguridad absoluta, mientras que en todas las situaciones intermedias habrá incertidumbre y está involucrado el cálculo de probabilidades.

De modo que obtenemos la primera conclusión general:

$$0 \leq p \leq 1$$

El símbolo $\leq$ se lee “menor o igual” y se establece así el rango permitido del número “probabilidad”. Los extremos son determinísticos y todo el rango intermedio $0 < p < 1$ es probabilístico.

Técnicas de conteo

Esta noción intuitiva de probabilidad conduce a cuestionarse cómo desarrollar técnicas que faciliten el conteo de casos posibles y favorables en un experimento dado. Existe un conjunto de técnicas más o menos ordenadas en términos de la generalidad de aplicación. No debe confundirse este primer problema con el objetivo último de la teoría de probabilidades, que consiste en el desarrollo de modelos y de elementos de juicio que faciliten la toma de decisión en situaciones de incertidumbre, problemas mucho más interesantes y complejos que los relativos a juegos de azar. Pero los juegos de azar tienen como característica que pretenden ofrecer una descripción completa y clara del procedimiento, y de las reglas del juego, de manera tal que se tiene un conocimiento total de los resultados posibles. Es un planteo que se aproxima mucho a la idealidad requerida para el desarrollo de cualquier teoría. Sin embargo, la realidad resulta mucho más compleja.

Planteamos en primera instancia un juego de azar en el que todos los resultados posibles son claramente distinguibles e igualmente posibles, o simétricos. Así es suficiente contar el número de caras de una moneda, de lados de un dado, de cartas en un mazo o de números en un bolillero. Pero la combinación de alternativas conduce a una progresiva complejidad que requiere un mínimo desarrollo teórico. En esta primera etapa nos limitaremos al conteo del número de casos posibles en un juego de azar.

Permutaciones

Uno de los primeros problemas de conteo es el de las permutaciones o número de ordenamientos lineales posibles, sin repetición, de un conjunto de elementos distinguibles que no tienen ninguna limitación acerca de la ubicación en el ordenamiento. El caso típico se ejemplifica en el conteo de la cantidad de maneras posibles en que puede ordenarse un número n de objetos distinguibles sin restricción en el ordenamiento permitido. Así un único objeto puede ordenarse de una única manera. Dos objetos de dos maneras (AB y BA), tres objetos de seis maneras (ABC, ACB, BAC, BCA, CAB y CBA). Si observamos la secuencia de casos posibles resulta 1, 2, 6, que coincide con 1 = 1, 2 = 2.1, 3 = 3.2.1 y la generalización inductiva consiste en el "factorial" de un número n de elementos distinguibles y simétricamente posible dado por $n!$. También puede plantearse, de modo más general, que el primer lugar puede ser ocupado por cualquiera de los n objetos, el segundo por cualquiera de los n-1 restantes, excepto el que fue ubicado en primer lugar, y así sucesivamente. Si razonamos entonces que por cada uno de los n posibles objetos ubicados en el primer lugar hay n-1 opciones para el segundo y que habrá n-2 para el tercero hasta una única opción para el último lugar, luego de haber ubicado a todos los demás, y además de interpreta el "por cada" como una multiplicación, se concluye que el número de permutaciones será $n!$. Esto conduce a una primera ecuación que se ofrece como técnica de conteo. Si P_n indica "permutaciones de n elementos distinguibles y simétricos", será

$$P_n = n!$$

Independencia y multiplicación

Tras la idea de multiplicar casos posibles se encubre la hipótesis de independencia entre ellos. Es decir que "por cada n posibilidades para el primer lugar hay n-1 para el segundo" presupone que el segundo lugar y los objetos que podrían ubicarse en el segundo lugar no se ven influenciados en manera alguna por el objeto ubicado en primer lugar. Esta "no influencia" de un resultado sobre otro es lo que se interpreta como independencia y es lo que se presupone en el cálculo del número de permutaciones. Como contra ejemplo puede plantearse ordenar diez fichas numeradas del 1 al 10. Si se plantea que la ficha que precede o antecede a otra no puede ser un número consecutivo, no habrá independencia entre una ficha y la precedente, es decir que, si por azar el primer lugar fue ocupado por la ficha 5, el segundo no puede ser ocupado por la 4 ni por la 6; pero entonces habrá 10 posibilidades para el primer lugar, aunque no habrá 9 para el segundo sino sólo 7 porque la 5 fue ubicada y la 4 y la 6 no están permitidas. Esto invalida el uso de $n!$ como técnica de conteo del número de permutaciones debido al condicionamiento impuesto en la secuencia.

En otras palabras, es la independencia mutua entre las distintas etapas de un experimento lo que nos permite interpretar el "por cada" como una simple multiplicación en sentido matemático. Para ello suele recurrirse gráficamente a la imagen de un árbol con sus ramificaciones.

Exclusión y adición

Un caso diferente se plantea si un jugador debe elegir sólo uno entre varios juegos y desea contabilizar el número de casos posibles de que dispone. Por ejemplo, puede elegir entre tirar una moneda, o tirar un dado, o sacar una carta de un mazo de 40, o extraer una bolilla entre 100. Como no puede realizar todas las experiencias en forma simultánea sino sólo una, tendrá 2 casos posibles para la moneda, 6 para el dado, 40 para la carta, 100 para el bolillero, por lo tanto, en total 2+6+40+100=148 casos posibles. Resulta aquí que la exclusión de partes de un experimento conduce a la adición de los casos posibles.

Variaciones

Una variación consiste en una limitación de la permutación a un número de elementos menor que el total disponible. Es decir que si de *n* elementos se pretende reordenar un subconjunto de $r < n$ objetos distinguibles y simétricamente posibles, se expresa el experimento con el término "variación"; simbólicamente nVr o "variación de *n* elementos tomados de a *r*". Puede observarse que $nVn = P_n$, es decir, una variación de *n* elementos tomados de a *n* coincide con una permutación de los *n* elementos.

Siguiendo el razonamiento anterior, se dispone de *n* posibilidades para el primer lugar, *n-1* para el segundo, *n-2* para el tercero, y así sucesivamente hasta *n-r+1* para el último lugar a ser ocupado. Este planteo puede expresarse como un cociente de factoriales.

$$nVr \ = \ n.(n-1).(n-2) \ldots\ldots (n-r+1) = \ \frac{n!}{(n-r)!}$$

La interpretación de este cociente puede plantearse en términos de eliminar del ordenamiento de todos los elementos ($n!$) el ordenamiento de aquellos que no se incorporan entre los *r* seleccionados $(n-r)!$

Observemos que hemos combinado el recurso a la multiplicación con el uso de la división. Hemos permutado todos los elementos para luego suprimir las permutaciones de los elementos que no integran la variación por medio de un cociente.

Combinaciones

Se llama de esta manera al conteo del número de agrupamientos posibles a partir de *n* elementos distinguibles y simétricos formando grupos de *r* elementos. El análisis puede plantearse de manera similar a los anteriores, pero quizá es más fácil partir de la idea de variación como agrupación de *n* elementos ordenados. Si se considera que en la variación se contabiliza el ordenamiento de los *r* elementos seleccionados como un caso posible diferente y de modo multiplicativo, es claro que suprimir el ordenamiento consiste en dividir por el número de ordenamientos en el subconjunto, es decir, por los ordenamientos internos del subconjunto que no se deben contabilizar como casos diferentes, esto es, dividir por $r!$.

Resulta así, si nCr se lee "combinaciones de *n* elementos tomados de a *r*" que

$$nCr \ = \ \frac{nVr}{r!} = \binom{n}{r} = \frac{n!}{(n-r)! \ r!}$$

y esto no es más que el número combinatorio. Otra vez hemos recurrido a contabilizar el número de permutaciones de todos los elementos (*n!*), dividir por las permutaciones de los elementos que no se incorporan a la combinación *(n-r)!*, y también dividir por las permutaciones internas de los elementos seleccionados dado que el ordenamiento interno no establece una diferenciación en cuáles elementos integran el grupo seleccionado *r!*.

Permutaciones con reposición

Con esta denominación puede referirse a la permutación de *n* elementos distinguibles y simétricos, pero que pueden aparecer repetidas veces en forma independiente dentro del ordenamiento. Por ejemplo, una secuencia de números decimales. Así un número de cinco cifras permite diez posibilidades para el primer lugar, diez para el segundo, tercero, cuarto y quinto, es decir, diez elevado a la quinta potencia. En general, para n casos distinguibles ubicados en r posiciones, resulta n^r (n a la r) ordenamientos posibles. La breve discusión inicial relativa a tirar dos monedas es un caso particular y el más simple de permutaciones con repetición.

Otras experiencias con juegos de azar

Si se dispone de un total de n objetos, pero no todos distinguibles de modo que hay r clases con n_i, $i = 1 \dots r$ elementos en cada clase indistinguibles entre sí, podemos construir una permutación de todos los elementos como si fuesen todos distinguibles y luego dividir el total por el producto de las permutaciones internas dentro de cada grupo.

$$\frac{n!}{n_1!\,n_2!\dots n_i!\dots n_r!} = \frac{(n_1 + n_2 + \cdots + n_i + \cdots + n_r)!}{n_1!\,n_2!\dots n_i!\dots n_r!}$$

Si por otra parte se extrae k objetos de un conjunto de n sin reposición, de manera tal que hay n_1 objetos del tipo 1, n_2 del tipo 2 y así sucesivamente hasta n_r objetos del tipo r, si se pregunta por el número de agrupaciones posibles tales que haya k_1 objetos del tipo 1, k_2 del tipo 2 y así siguiendo, habrá

$$\binom{n_1}{k_1}\binom{n_2}{k_2}\dots\binom{n_i}{k_i}\dots\binom{n_r}{k_r}$$

grupos diferentes que pueden formarse. Por ejemplo, si se forman comisiones de 1 alumno de quinto, con 10 estudiantes, 2 de cuarto con 15, 3 de tercero con 20, 4 de segundo con 25 y 5 de primero con 30 resulta que habrá ${}_{10}C_1 * {}_{15}C_2 * {}_{20}C3 * {}_{25}C_4 * {}_{30}C_5$ comisiones diferentes a formar.

Definición clásica o empírica de probabilidad

Por lo discutido hasta aquí, es claro que el problema de la incertidumbre condujo a la búsqueda de respuestas en el ámbito de las matemáticas. Estas respuestas tenían por objetivo resolver un problema concreto y no confeccionar un desarrollo teórico. Sin embargo el planteo matemático requiere de un marco teórico formal. Por este motivo se suele distinguir entre la "definición" empírica de probabilidad, que en realidad no es una definición sino una herramienta que parte del cálculo para contribuir a la solución de un problema concreto y tampoco es empírica sino hipotética suponiendo que se cumplen las reglas del juego, de la definición formal o matemática, que trataremos más adelante.

Una primera definición ya fue esbozada de manera intuitiva, ahora llamada "empírica", por medio de un cociente entre el número de casos favorables y el número de casos posibles. Así

$$p = \frac{cf}{cp} = \frac{número\ de\ casos\ favorables}{número\ de\ casos\ posibles}$$

es una definición que en primera aproximación se ajusta a la idea intuitiva de probabilidad. Este número estará en el rango entre 0 y 1. Vale 0 cuando no hay casos favorables, por lo que se está absolutamente seguro del fracaso, y 1 cuando todos los casos posibles son favorables y, por lo tanto, se está absolutamente seguro del éxito. Ambos extremos pueden considerarse como "determinísticos" debido a la seguridad absoluta acerca del resultado. Es claro que en tal caso no hay ninguna decisión que tomar sino una "determinación". También es claro que la equiprobabilidad es la peor situación en que puede encontrarse una persona que pretende usar el cálculo de probabilidades para tomar una decisión porque este número le dirá que todas las decisiones son igualmente probables y que por lo tanto todas entrañan el mismo nivel de riesgo. Esta es la condición de los juegos de azar.

Introducción al desarrollo histórico

En este marco de ideas se desarrolló la primera etapa de la teoría de probabilidades, limitada a los juegos de azar y haciendo uso de las herramientas matemáticas como técnicas de cálculo.

Tras las comunicaciones entre Pascal y Fermat, en el siglo XVII, sobre los juegos de azar, en manos de James Bernoulli se publica en 1713 por primera vez un teorema de validez general y Abraham de Moivre publica otro teorema, conocido luego como “de la multiplicación” en 1718, y en 1738 otro teorema que conducirá más tarde a la “distribución normal”.

Comienza a verse la conexión entre esta teoría matemática de los juegos de azar y los eventos relativos a poblaciones humanas (nacimientos, defunciones). Esta extensión del campo de la teoría de probabilidades dio buenos resultados en relación con expectativas sobre las rentas, seguros de vida y otras aplicaciones similares. Sin embargo, no se tuvo en cuenta aún que la definición empírica presupone el conocimiento de los casos favorables y posibles, lo cual sólo es válido en los juegos de azar, pero no en la evaluación de rentas. En 1812, Laplace expone una síntesis de los conocimientos relativos a la teoría de probabilidades, tanto aplicada a los juegos de azar como a otras cuestiones científicas y técnicas. Pero no cuestiona la validez de la definición clásica o empírica, y la admite como aplicable a todos los casos. Desde entonces se amplió el campo de aplicaciones de la teoría de probabilidades. Gauss y Laplace desarrollaron una teoría sistemática de errores de medición, hacia mediados del siglo XIX se aplicó a la física desarrollando la teoría cinética de los gases, luego termodinámica estadística y mecánica estadística en el siglo XX. El campo actual de aplicaciones y nivel de desarrollo la ha convertido en una de las herramientas básicas en casi cualquier campo de las ciencias y la tecnología.

Sin embargo, el desarrollo de las aplicaciones no fue acompañado por la consolidación de los fundamentos teóricos, persistiendo la validez de la definición clásica, empírica e intuitiva, aunque estaba desde el principio limitada a los juegos de azar. Este desarrollo se llevó a cabo desde mediados del siglo XIX y aún persiste en progresiva consolidación a lo largo del siglo XX y comienzos del XXI. La definición clásica ha sido reemplazada por definiciones que hacen uso de razones frecuenciales y, desde el punto de vista formal, por una definición axiomática.

Formulación frecuencial de la noción de probabilidad

La “definición” clásica se relaciona con un principio empírico que parte de la repetición de un mismo experimento un gran número de veces. La experiencia indica que en tal caso la frecuencia de éxitos es proporcional a la razón entre casos exitosos dentro del conjunto de casos posibles. Es decir que, en un dado, si se pretende estimar experimentalmente la probabilidad de que se obtenga un seis, como hay seis casos posibles y sólo uno favorable, la definición empírica indica que p = 1/6, y la experiencia mostrará que en 6 tiradas podría obtenerse un seis, pero quizá ninguno, tal vez dos, aunque rara vez todos serán seis. Si se tira 60 veces, se obtendrá un número en torno a 10 de veces en las que se obtiene un seis. Será raro que no salga ninguna vez y extremadamente raro que siempre salga seis. En 600 tiradas se obtendrá del orden de cien veces seis, en 6000 tiradas, del orden de mil veces seis. En todos los casos la proporción de éxitos a repeticiones será del orden de 1/6, y tanto más se aproximará a este valor cuanto más grande sea el número de repeticiones.

De esto se obtiene que, si se llama f al número de éxitos o frecuencia de éxitos y n al número de repeticiones del experimento en igualdad de condiciones, será

$$\lim_{n\to\infty} \frac{f}{n} = p$$

Es decir que cuando n tiende a infinito, la frecuencia relativa de éxitos a repeticiones tenderá a la probabilidad real o experimental, que debería coincidir con la probabilidad empírica o hipotética si se cumplen las reglas del juego. Sin embargo, puede notarse que esta definición que aplica la noción de límite de sucesiones, que también puede llamarse experimental o estadística de la probabilidad, no requiere que se limite a juegos de azar y puede extenderse a situaciones reales en tanto exista uniformidad de las condiciones en las que se realizó el experimento. En tales casos conocido el hecho que aun así hay una variabilidad residual o intrínseca en todas las etapas de la realización del experimento que conduce a resultados diferentes. Estas fuentes de variabilidad están fuera del control del experimentador a pesar de todos los recaudos que pretenda tomar.

La aplicación particular a una medición nos lleva a considerar que esencialmente ésta consiste, en primera instancia, en la interacción entre un objeto a medir, un instrumento de medición y un patrón de referencia que define una unidad. El procedimiento de medición exige precisar el método a seguir para que esta interacción dé por resultado una medición de la magnitud en cuestión. Precisamente la descripción detallada, precisa y completa del procedimiento es un experimento aleatorio. Se desconoce el resultado de la medición antes de ser realizada y persiste una variabilidad intrínseca que provoca que los resultados de sucesivas mediciones, aun en las mismas condiciones de realización, sean diferentes. Esto conduce a pensar que el cálculo de probabilidades puede contribuir a una teoría sobre los errores de medición.

Pero para ello se requiere algo más. Si sucesivas mediciones dan resultados diferentes, ¿cuál debe tomarse como válida? Es aquí cuando se tiene en cuenta, según veremos más adelante, que un gran número de mediciones define un valor más confiable. Esta observación sobre el gran número de mediciones se apoya en la definición de la probabilidad sobre la base de una razón frecuencial.

Probabilidades y Estadística: Ejercicios Unidad Nº 1

Generalidades sobre Combinatoria y Cálculo de probabilidades

1. ¿De cuántas maneras se pueden colocar 10 libros en un estante?, ¿de cuántas formas si 4 de ellos determinados deben ocupar los mismos lugares aun cuando puedan intercalarse entre sí?

2. ¿De cuántas maneras se pueden colocar en fila 6 hombres si uno determinado no puede estar nunca a la cabeza?

3. ¿De cuántas maneras se pueden seleccionar y ordenar subconjuntos de 4 dígitos diferentes de los números entre el 0 y el 9? ¿De cuántas maneras si los dígitos entre el 0 y el 9 pueden repetirse dentro de la misma cifra?

4. ¿Cuántos triángulos se pueden formar con diez puntos no alineados en un plano?

5. ¿Cuántos caracteres se pueden formar con los dos dígitos binarios si se configuran cadenas en las cuales entran hasta 8 unos y ceros?

6. Una urna contiene 10 bolillas numeradas del 0 al 9. Se sacan sucesivamente al azar 5 bolillas (con reposición). Se busca la probabilidad de que juntando los números de cada una en el orden de extracción resulte el número 21345. Hallar esta probabilidad si las bolillas se extraen sin reposición.

7. Se arrojan sucesivamente dos dados, calcular la probabilidad de que el primero salga seis, de que el segundo sea par, y de que ocurran ambas cosas simultáneamente.

8. Se tiene 5 pares de zapatos mezclados y cada par es distinto de los demás. Si se eligen 2 zapatos al azar, ¿hallar la probabilidad de que correspondan a un mismo par?

9. Se ordenan 10 libros en un estante, calcular el número de casos posibles de ordenamientos. Si se pretende que dos de ellos determinados queden juntos en el ordenamiento aleatorio, calcular el número de casos favorables y la probabilidad de éxito. Si además se pretende que los dos libros determinados queden ordenados por tomos, calcular el número de casos favorables y la probabilidad de éxito.

10. Se encuentran diez libros en un estante superior y pretende reordenar cinco de esos libros en un estante inferior. Hallar la probabilidad de que dos de los cinco libros extraídos y previamente determinados, queden juntos en el ordenamiento aleatorio.

11. Un agente de bienes raíces muestra casas a un comprador potencial. Hay diez casas del precio deseado en una lista de la zona. El comprador tiene tiempo para visitar sólo tres de ellas.
    a. ¿En cuántas formas podrían escogerse las tres casas, si se considera el orden de visita?
    b. ¿En cuántas formas podrían escogerse las tres casas, si el orden no es importante?
    c. Si cuatro de las casas son nuevas y seis han sido ocupadas previamente, y las tres casas por visitar se escogen al azar, ¿cuál es la probabilidad de que sean nuevas?

12. Al poco tiempo de ser puestos en servicio, algunos autobuses fabricados por cierta compañía presentan grietas en la parte inferior del bastidor principal; suponga que una ciudad tiene 20 de estos autobuses y que han aparecido grietas en 8 de ellos.
    a. ¿Cuántas formas hay de seleccionar una muestra de 5 autobuses de los 20 para una inspección completa?

b. ¿En cuántas formas puede una muestra de 5 autobuses contener exactamente 4 con grietas visibles?
c. Si se escoge al azar una muestra de 5 autobuses, ¿cuál es la probabilidad de que exactamente 4 de los 5 tengan grietas visibles?
d. Si se seleccionan los autobuses como en “c”, ¿cuál es la probabilidad de que al menos 4 de los seleccionados tengan grietas visibles?

13. Un investigador estudia los efectos de la temperatura, presión y el tipo de catalizador de cierta reacción química. Considera tres diferentes temperaturas, cuatro diferentes presiones y cinco diferentes catalizadores.
a. Si en cualquier corrida experimental interviene el uso de una sola temperatura, presión y catalizador, ¿cuántas corridas experimentales son posibles?
b. ¿En cuántas corridas experimentales interviene el uso de la temperatura y las dos presiones más bajas?

14. Una planta de producción emplea 20 trabajadores en el turno de día, 15 en el segundo turno y 10 en el turno de la noche. Un consultor de control de calidad selecciona seis de estos trabajadores para hacerles una entrevista a fondo. Supongamos que la selección se hace en tal forma que cualquier grupo de seis trabajadores tiene la misma posibilidad de ser seleccionado, del mismo modo que cualquier otro grupo (selección de 6 sin sustitución entre 45).
A. ¿De cuántas maneras se pueden seleccionar seis trabajadores que provengan del turno de día? ¿Cuál es la probabilidad de que los seis trabajadores seleccionados sean del turno de día?
B. ¿Cuál es la probabilidad de que los seis trabajadores seleccionados sean del mismo turno?
C. ¿Cuál es la probabilidad de que al menos dos turnos diferentes sean representados entre los trabajadores seleccionados?

15. En cierta bodega una caja contiene cuatro focos de 40W, cinco de 60W y seis de 75W. Suponga que se seleccionan al azar tres focos.
A. ¿Cuál es la probabilidad de que exactamente dos de los focos elegidos sean de 75W?
B. ¿Cuál es la probabilidad de que los tres focos seleccionados tengan la misma potencia?
C. ¿Cuál es la probabilidad de seleccionar un foco de cada potencia?

16. Tres moléculas tipo A, tres B, tres C y tres D deben combinarse para formar una molécula en cadena. Una de estas moléculas por ejemplo en cadena es ABCDABCDABCD y otra es DCDDABABDBBCC.
A. ¿Cuántas de estas moléculas en cadena hay? (Sugerencia: si las tres A pudieran distinguirse una de otra A1, A2 y A3, y también las B, C y D, ¿Cuántas moléculas habría? ¿Cómo se reduce este número cuando los subíndices se eliminan de las A?)
B. Suponga que una molécula del tipo descrito es seleccionada al azar. ¿Cuál es la probabilidad de que las tres moléculas de cada tipo terminen una junto a otra (por ejemplo en BBBAAADDDCCC)?

Unidad 1. Ejercicios resueltos

En lo que sigue se resolverán los ejercicios de la unidad 1. Se espera que no se recurra a la resolución sino después de haberlos resuelto en forma personal o al menos el haberlo intentado. También sugerimos que al finalizar cada ejercicio se trate de responder al menos tres preguntas ¿cómo se vincula el ejercicio con la teoría?, ¿qué se aprendió con el ejercicio?, y ¿qué más se puede aprender introduciendo variantes o preguntas adicionales?

1. ¿De cuántas maneras se pueden colocar 10 libros en un estante?, ¿de cuántas formas si 4 de ellos determinados deben ocupar los mismos lugares aun cuando puedan intercalarse entre sí?

Si consideramos que los libros son distinguibles, a menos que sea una librería o una biblioteca con libros iguales, la ubicación ordenada corresponde a una aplicación inmediata de la técnica de permutaciones. De modo que N=10! es la respuesta a la primera pregunta.
Cuando se plantea que cuatro de ellos tienen lugares preestablecidos, es claro que también los otros seis libros en los seis lugares restantes. Dentro de esos lugares pueden ser ordenados sin otra restricción en 4!=24 maneras, mientras que los otros seis libros serán ordenados de 6!=720 modos diferentes posibles.
El número total de ordenamiento de los diez libros con esta restricción es el producto de las dos posibilidades. Podemos pensarlo, de acuerdo con el principio de multiplicación, como “por cada una de las 24 maneras de ordenar los cuatro en sus lugares hay 720 maneras de ordenar los otros seis en sus lugares respectivos, por lo que habrá un total de

$$N = 4! * 6! = 24 * 720 = 17180$$

Ordenamientos diferentes frente a los 10!=3628800 permutaciones posibles sin restricciones.

2. ¿De cuántas maneras se pueden colocar en fila 6 hombres si uno determinado no puede estar nunca a la cabeza?

Si no hubiese ninguna restricción, la respuesta sería la correspondiente a la permutación de seis elementos distinguibles, por lo tanto 6!=720 ordenamientos. Pero como una de las personas no puede estar delante, basta restar ese caso, que se contabiliza ubicando a la persona no permitida en el primer lugar y contar de cuántas maneras se pueden ubicar las otras cinco detrás, lo que da 5!=120 formas diferentes. Luego tenemos un total de N=6!-5!=720-120=600 ordenamientos posibles.
Otra manera de abordarlos es separar el primer lugar y contabilizar las cinco personas permitidas para ocuparlo, por cada una de ellas debemos multiplicar luego por las 5! maneras diferentes de ordenar a las cinco personas que no tienen restricciones.

$$N = 6! - 5! = 5 * 5! = 600$$

Si hacemos un simple cálculo, veremos que ambas expresiones conducen al mismo resultado. Una conclusión inmediata es observar que un mismo problema admite más de un abordaje para llegar a la solución.

3. ¿De cuántas maneras se pueden seleccionar y ordenar subconjuntos de 4 dígitos diferentes de los números entre el 0 y el 9? ¿De cuántas maneras si los dígitos entre el 0 y el 9 pueden repetirse dentro de la misma cifra?

La primera pregunta describe el caso paradigmático de variaciones: seleccionar y ordenar subconjuntos. La respuesta es inmediata

$$N = V_{10}^{4} = \frac{10!}{(10-4)!} = 5040$$

La segunda pregunta responde al caso típico de permutaciones con repetición, ubicar en cuatro lugares uno cualquiera de diez elementos tantas veces como se desee. Nos da

$$N = 10 * 10 * 10 * 10 = 10^4 = 10000$$

permutaciones diferentes con repetición. Es equivalente a preguntar cuántos números de cuatro cifras se pueden formar en el sistema decimal.

4. ¿Cuántos triángulos se pueden formar con diez puntos no alineados en un plano?

Para responder a este ejercicio se puede ayudar visualmente con el dibujo de diez puntos no alineados y dos o tres triángulos entre tres puntos cualesquiera. No importa el orden en que se elijan esos puntos, siempre se formará el mismo triángulo en tanto sean los mismos tres puntos. Esto nos lleva a pensar en que el problema es la selección de puntos y no el ordenamiento, lo que coincide con una combinación, de modo que

$$N = \binom{10}{3} = \frac{10!}{3!\,(10-3)!} = 120$$

triángulos.

5. ¿Cuántos caracteres se pueden formar con los dos dígitos binarios si se configuran cadenas en las cuales entran hasta 8 unos y ceros?

Si se forma un número binario con ocho caracteres, dado que corresponde a una permutación con repetición, el número total será 2^8=256 secuencias binarias diferentes. (Corresponde al número de caracteres del código ASCII)
Si observamos que dice “hasta” ocho unos y ceros, la palabra “hasta” nos dice que puede haber menos de ocho caracteres binarios, es decir que se suman las potencias menores que ocho al número anterior.

$$N = 2^8 + 2^7 + 2^6 + 2^5 + 2^4 + 2^3 + 2^2 + 2^1 = 510$$

El motivo por el cual se suman es que aplicamos el principio de exclusión. Si se decide armar secuencias de seis caracteres binarios, por ejemplo, no se podrá seleccionar una secuencia de cinco ni de siete ni de ninguna otra longitud. Elegir el número de caracteres o longitud de la cadena excluye la selección de otras cantidades de caracteres. Por lo tanto hay dos etapas involucradas, primero la selección de la longitud de la cadena y, para cada selección excluyendo a las otras, habrá la potencia de dos correspondiente a la longitud como número de cadenas posibles.

6. Una urna contiene 10 bolillas numeradas del 0 al 9. Se sacan sucesivamente al azar 5 bolillas (con reposición). Se busca la probabilidad de que juntando los números de cada una en el orden de extracción resulte el número 21345. Hallar esta probabilidad si las bolillas se extraen sin reposición.

Este ejercicio introduce la noción de probabilidad. Al sacarlas sucesivamente con reposición, siempre hay diez bolillas y el número de casos posibles corresponde a una permutación con reposición. Si la bolilla no vuelve a reponerse en el bolillero, la secuencia será de cinco bolillas diferentes, con lo cual corresponde a una variación. En ambos casos hay solamente un caso favorable, de modo que la probabilidad de la secuencia 21345 vale

$$p(21345) = \frac{1}{10^5} \qquad y \qquad p(21345) = \frac{1}{V_{10}^5} = \frac{1}{30240}$$

7. Se arrojan sucesivamente dos dados, calcular la probabilidad de que el primero salga seis, de que el segundo sea par, y de que ocurran ambas cosas simultáneamente.

En primer lugar calculemos el número de resultados posibles de tirar dos dados. Corresponde a una permutación con repetición teniendo seis resultados posibles y dos repeticiones. Visualmente podemos verlo como los pares ordenados 11, 12, 13, 14, 15, 16, 21, 22,……,64,65,66. Usando permutaciones con repetición tenemos $N=6^2=36$ casos posibles.
Si los escribimos como pares, es fácil observar que seis de esos pares comienzan con el número 6, de modo que

$$p(6\,*) = \frac{6}{36} = \frac{1}{6} = 0{,}1\hat{6} = 16{,}\hat{6}\%$$

El número de casos en el que el segundo es para vale 18, la mitad del total, luego

$$p(*\,par) = \frac{18}{36} = \frac{1}{2} = 0{,}5 = 50\%$$

En sólo tres de los que comienzan con el número 6 el segundo es par, son los pares ordenados 62, 64 y 66. Luego

$$p(6 * y * par) = \frac{3}{36} = \frac{1}{12} = 0{,}08\hat{3} = 8{,}\hat{3}\%$$

Si bien es sencillo como ejercicio de conteo, nos enfrenta con el manejo de dos dados y, lo que es más importante, con observar que la conjunción lógica de pedir dos condiciones simultáneas nos llevó a un resultado equivalente al producto de las probabilidades.

8. Se tiene 5 pares de zapatos mezclados y cada par es distinto de los demás. Si se eligen 2 zapatos al azar, ¿hallar la probabilidad de que correspondan a un mismo par?

Una visión rápida de este ejercicio nos lleva a observar que no tiene relevancia cuál sea el primer zapato seleccionado, lo que se considera como favorable es que el segundo zapato sea el par correspondiente al primero. Después de haber sacado un zapato quedan nueve y sólo uno es el par del primero, por lo tanto la respuesta es inmediata y da 1/9.
Sin embargo podemos intentar un planteo más sofisticado y que aplique las técnicas de conteo. Como casos posibles tenemos la selección de dos zapatos sobre un total de diez, por lo tanto es el número de combinaciones de diez zapatos tomados en grupos de dos

$$cp = C_{10}^2 = \binom{10}{2} = 45$$

El número de casos favorables lo podemos plantear como la selección de un par de los cinco disponibles, formalmente

$$cf = C_5^1 = \binom{5}{1} = 5$$

El cociente entre casos favorables y posibles da nuevamente 1/9.

9. Se ordenan 10 libros en un estante, calcular el número de casos posibles de ordenamientos. Si se pretende que dos de ellos determinados queden juntos en el ordenamiento aleatorio, calcular el número de casos favorables y la probabilidad de éxito. Si además se pretende que

los dos libros determinados queden ordenados por tomos, calcular el número de casos favorables y la probabilidad de éxito.

El número de ordenamientos posibles de diez libros ya es inmediato, corresponde a las permutaciones de diez elementos o cp=10!. Para plantear, como caso favorable, que dos de ellos queden juntos, puede recurrirse a idealmente "atar" los libros y pensar que es un solo libro que se ordenan con los ocho restantes, por lo tanto nos daría 9! Ordenamientos diferentes de nueve libros. Pero a su vez estos dos libros se pueden permutar internamente, de modo que tenemos 2! permutaciones de los libros que siguen estando juntos en el orden invertido. Luego el número de casos favorables será el producto de las permutaciones de los nueve por las permutaciones de los dos, es decir cf=9!*2!, y la probabilidad de que en un ordenamiento aleatorio esto ocurra por azar vale

$$p = \frac{cf}{cp} = \frac{9! * 2!}{10!} = 0{,}2 = 20\%$$

Si los dos libros deben estar en un orden determinado, basta suprimir del 2! del planteo anterior y tendremos una probabilidad de un diez por ciento.
También podemos pensarlo considerando que los dos libros forman uno pero son permutables, luego tenemos 2!=2 ordenamientos posibles. Los otros ocho libros se pueden ordenar de cualquier manera, con lo cual tenemos 8! ordenamientos posibles de los ocho restantes. Pero luego ese par de libros unidos se puede ubicar en cualquiera de los dos extremos o en siete posiciones intermedias, por lo tanto por cada ordenamiento de los dos y a la vez de los ocho, tenemos nueve ubicaciones del par entre los libros restantes, por lo tanto

$$cf = 2! * 8! * 9$$

Es fácil verificar que el resultado es el mismo que el anterior.
Más allá de ser alternativas diferentes, nos ofrece una metodología. Primero comenzar por el cálculo de los casos posibles por ser menos restrictivo y en general más sencillo de resolver. Luego imponer la restricción y contabilizar el número de formas residuales de casos favorables dado que la restricción fue impuesta, en nuestro caso mantener unidos los dos libros. Finalmente combinar la libertad interna de la restricción, en nuestro ejemplo la permutación de los dos, con la libertad residual, la permutación de los ocho, en todas las maneras posibles, que en nuestro caso fue multiplicar por nueve.

10. Se encuentran diez libros en un estante superior y pretende reordenar cinco de esos libros en un estante inferior. Hallar la probabilidad de que dos de los cinco libros extraídos y previamente determinados, queden juntos en el ordenamiento aleatorio.

Considerando el orden de razonamiento propuesto a través del ejercicio anterior, primero calculemos el número de casos posibles. Dado que hay que ordenar un subconjunto de cinco elementos sobre un total de diez, se trata de una variación de diez elementos tomados en grupos de cinco elementos, lo que da

$$cp = V_{10}^{5} = \frac{10!}{(10-5)!} = 30240$$

Para calcular el número de casos favorables, impongamos la condición "favorable" y veamos el número de ordenamientos libres que permite. Tomamos los dos libros y los ubicamos en dos lugares del otro estante. Quedan ocho libros y tres estantes vacíos. No hay restricción sobre ellos, de modo que hay una variación de ocho elementos tomados en grupos de tres como ordenamientos libres. Pero los dos libros considerados pueden permutarse 2! veces. También pueden ubicarse juntos en los extremos o en dos lugares intermedios, de modo que hay cuatro formas diferentes de presentar las opciones anteriores, luego

$$cf = V_8^3 * 2! * 4 = \frac{8!}{(8-3)!} * 2! * 4 = 2688$$

Finalmente

$$p = \frac{cf}{cp} = \frac{2688}{30240} = 0{,}0\hat{8} \cong 8{,}9\%$$

11. Un agente de bienes raíces muestra casas a un comprador potencial. Hay diez casas del precio deseado en una lista de la zona. El comprador tiene tiempo para visitar sólo tres de ellas.
    a. ¿En cuántas formas podrían escogerse las tres casas, si se considera el orden de visita?
    b. ¿En cuántas formas podrían escogerse las tres casas, si el orden no es importante?
    c. Si cuatro de las casas son nuevas y seis han sido ocupadas previamente, y las tres casas por visitar se escogen al azar, ¿cuál es la probabilidad de que sean nuevas?

El punto “a” es claramente una variación de diez elementos tomados de a tres. En el caso “b”, si el orden no es importante, es una combinación. En el caso “c” se plantea que hay cuatro casas nuevas y seis previamente ocupadas. El número de casos favorables de obtener al azar tres casas nuevas entre las cuatro equivale a, si se las selecciona en forma ordenada, una variación de cuatro tomando grupos de tres casas, y, si no importa el orden sería una combinación de tres casas seleccionadas entre cuatro. En un caso calcularíamos $V_4^3 = \frac{4!}{(4-3)!} = 24$ formas diferentes, y en el otro sería $C_4^3 = 4$ modos distintos. Si realizamos los cálculos en las dos situaciones, si importa el orden

$$p = \frac{V_4^3}{V_{10}^3} = \frac{24}{720} = 0.0\hat{3} \qquad y \qquad p = \frac{C_4^3}{C_{10}^3} = \frac{4}{120} = 0.0\hat{3}$$

En ambos casos se obtiene el mismo resultado porque la permutación del orden interno del grupo seleccionado se omite tanto del numerador como del denominador en la combinación.

12. Al poco tiempo de ser puestos en servicio, algunos autobuses fabricados por cierta compañía presentan grietas en la parte inferior del bastidor principal; suponga que una ciudad tiene 20 de estos autobuses y que han aparecido grietas en 8 de ellos.
    a. ¿Cuántas formas hay de seleccionar una muestra de 5 autobuses de los 20 para una inspección completa?
    b. ¿En cuántas formas puede una muestra de 5 autobuses contener exactamente 4 con grietas visibles?
    c. Si se escoge al azar una muestra de 5 autobuses, ¿cuál es la probabilidad de que exactamente 4 de los 5 tengan grietas visibles?
    d. Si se seleccionan los autobuses como en “c”, ¿cuál es la probabilidad de que al menos 4 de los seleccionados tengan grietas visibles?

El punto “a” es claramente una combinación de veinte elementos tomados en grupos de cinco. El punto “b” propone en cambio que una selección de cuatro elementos sea tomada de los ocho con grietas visibles, de modo que es una combinación de cuatro elementos tomados a partir de ocho. Sin embargo notemos que por cada combinación de cuatro autobuses con grietas visibles hay un quinto elemento que puede ser cualquiera de los doce con grietas ocultas. De modo que tendremos

$$N = C_8^4 C_{12}^1 = \binom{8}{4}\binom{12}{1} = 70 * 12 = 840$$

En el punto “c” se pide el cálculo de la probabilidad de que la selección propuesta en el punto “b” se obtenga por azar, por lo tanto un cociente entre los dos resultados anteriores.

$$p = \frac{cf}{cp} = \frac{\binom{8}{4}\binom{12}{1}}{\binom{20}{5}} = \frac{\binom{8}{4}\binom{12}{1}}{\binom{8+12}{4+1}}$$

La última expresión es un caso particular de un desarrollo más general, que se conoce como "distribución hipergeométrica", que estudiaremos más adelante y es útil en la toma de muestras sobre poblaciones pequeñas.
El punto "d" propone algo similar al punto "c", pero al decir "al menos cuatro" incluye la probabilidad de que sean 4 o 5 los autobuses con grietas visibles. Como ambas posibilidades son excluyentes, conduce a la suma del número de opciones y, dado que es el mismo denominador, a la suma de probabilidades.

$$p = \frac{cf}{cp} = \frac{\binom{8}{4}\binom{12}{1}}{\binom{8+12}{4+1}} + \frac{\binom{8}{5}\binom{12}{0}}{\binom{8+12}{5+0}}$$

13. Un investigador estudia los efectos de la temperatura, presión y el tipo de catalizador de cierta reacción química. Considera tres diferentes temperaturas, cuatro diferentes presiones y cinco diferentes catalizadores.
    a. Si en cualquier corrida experimental interviene el uso de una sola temperatura, presión y catalizador, ¿cuántas corridas experimentales son posibles?
    b. ¿En cuántas corridas experimentales interviene el uso de la temperatura y las dos presiones más bajas?

Este problema es simple y sólo refuerza las nociones de independencia en el conteo. A la pregunta (a) debemos responder con el producto de las tres temperaturas, cuatro presiones y cinco catalizadores obteniendo 60 combinaciones diferentes. A la pregunta (b) nos limitamos usando la temperatura (1) más baja, y las dos (2) presiones menores combinadas con los cinco catalizadores obteniendo 10 opciones.

14. Una planta de producción emplea 20 trabajadores en el turno de día, 15 en el segundo turno y 10 en el turno de la noche. Un consultor de control de calidad selecciona seis de estos trabajadores para hacerles una entrevista a fondo. Supongamos que la selección se hace en tal forma que cualquier grupo de seis trabajadores tiene la misma posibilidad de ser seleccionado, del mismo modo que cualquier otro grupo (selección de 6 sin sustitución entre 45).
    A. ¿De cuántas maneras se pueden seleccionar seis trabajadores que provengan del turno de día? ¿Cuál es la probabilidad de que los seis trabajadores seleccionados sean del turno de día?
    B. ¿Cuál es la probabilidad de que los seis trabajadores seleccionados sean del mismo turno?
    C. ¿Cuál es la probabilidad de que al menos dos turnos diferentes sean representados entre los trabajadores seleccionados?

El problema 14 es una aplicación de la distribución hipergeométrica. El número de casos posibles será el combinatorio de seis elementos tomados entre un grupo de 45 distinguibles.
Para responder a la pregunta (a) debemos considerar las combinaciones en las que pueden tomarse los 6 trabajadores entre el grupo de 20 del turno de día. En términos de cálculo

$$p = \frac{\binom{20}{6}}{\binom{45}{6}} = \frac{\binom{20}{6}\binom{15}{0}\binom{10}{0}}{\binom{20+15+10}{6+0+0}}$$

Es fácil extender el razonamiento anterior al considerar las tres opciones en que pueden provenir del mismo turno.

$$p = \frac{\binom{20}{6}\binom{15}{0}\binom{10}{0}}{\binom{20+15+10}{6+0+0}} + \frac{\binom{20}{0}\binom{15}{6}\binom{10}{0}}{\binom{20+15+10}{6+0+0}} + \frac{\binom{20}{0}\binom{15}{0}\binom{10}{6}}{\binom{20+15+10}{6+0+0}}$$

En el planteo (c) pedimos que no provengan los tres del mismo turno, por lo tanto el planteo es complementario al del punto (b) y basta restar ésta probabilidad a 1.

15. En cierta bodega una caja contiene cuatro focos de 40W, cinco de 60W y seis de 75W. Suponga que se seleccionan al azar tres focos.
    A. ¿Cuál es la probabilidad de que exactamente dos de los focos elegidos sean de 75W?
    B. ¿Cuál es la probabilidad de que los tres focos seleccionados tengan la misma potencia?
    C. ¿Cuál es la probabilidad de seleccionar un foco de cada potencia?

Consecuentemente con el razonamiento del problema anterior, tenemos, para el punto (a)

$$p = \frac{\binom{4}{1}\binom{5}{0}\binom{6}{2}}{\binom{4+5+6}{1+0+2}} + \frac{\binom{4}{0}\binom{5}{1}\binom{6}{2}}{\binom{4+5+6}{0+1+2}}$$

Para el punto (b)

$$p = \frac{\binom{4}{3}\binom{5}{0}\binom{6}{0}}{\binom{4+5+6}{3+0+0}} + \frac{\binom{4}{0}\binom{5}{3}\binom{6}{0}}{\binom{4+5+6}{0+3+0}} + \frac{\binom{4}{0}\binom{5}{0}\binom{6}{3}}{\binom{4+5+6}{0+0+3}}$$

Y para el (c)

$$p = \frac{\binom{4}{1}\binom{5}{1}\binom{6}{1}}{\binom{4+5+6}{1+1+1}}$$

16. Tres moléculas tipo A, tres B, tres C y tres D deben combinarse para formar una molécula en cadena. Una de estas moléculas por ejemplo en cadena es ABCDABCDABCD y otra es DCDDABABDBBCC.
    A. ¿Cuántas de estas moléculas en cadena hay? (Sugerencia: si las tres A pudieran distinguirse una de otra A1, A2 y A3, y también las B, C y D, ¿Cuántas moléculas habría? ¿Cómo se reduce este número cuando los subíndices se eliminan de las A?)
    B. Suponga que una molécula del tipo descrito es seleccionada al azar. ¿Cuál es la probabilidad de que las tres moléculas de cada tipo terminen una junto a otra (por ejemplo en BBBAAADDDCCC)?

Si fueran todas distinguibles, tendríamos doce moléculas diferente con 12! permutaciones distintas. Como dentro de cada tipo hay tres moléculas idénticas, sus permutaciones internas son indistinguibles. Dado que hay 3! permutaciones dentro de cada tipo, debemos dividir las permutaciones supuestas distinguibles por las permutaciones internas dentro de cada grupo contabilizadas multiplicativamente en exceso, luego, para el punto (a)

$$N = \frac{12!}{3! * 3! * 3! * 3!} = \frac{12!}{(3!)^4} = 369600$$

Para el punto (b), como cada grupo forma una unidad, habría 4! permutaciones de los grupos configurados de esa manera, luego

$$p = \frac{4!}{12!/(3!)^4} = 0{,}0065\%$$

Unidad 2

Cálculo Elemental de Probabilidades y aplicaciones de teoremas básicos

Definición axiomática de probabilidad

La definición experimental, que recurre a la noción de límite, tiene la desventaja que no se trata de una definición formal y que requiere la aceptación "a priori" de la validez de la tendencia a la estabilidad de la razón de frecuencias ante un número progresivamente más grande de repeticiones (si bien este planteo se demostrará más adelante en un teorema).

En un marco más formal, partimos de la descripción del experimento aleatorio E. Se llamará p a una función del experimento E y del resultado posible A considerado como exitoso para ese experimento. Notamos

$$p = P(E, A)$$

Se admite que existe un valor real o verdadero al que la experimentación sólo permite aproximarse por medio de la razón de frecuencias. Cuando el experimento se presupone conocido, se omite de la expresión y escribe $p = p(A)$ como probabilidad únicamente asociada al resultado A, habitualmente llamado "evento" o "suceso" A.

Para los juegos de azar este valor p puede calcularse a priori asumiendo que los resultados posibles son simétricos, distinguibles y totalmente conocidos. En otras ocasiones sólo se dispone de una aproximación a través de una razón de frecuencias experimental. De esta manera se extiende la definición de probabilidad a un número más amplio de situaciones y aplicaciones, más próximas al mundo real que la simple aplicación a los juegos de azar.

Pueden proponerse tres axiomas (una teoría axiomática completa requiere además la definición de suceso incorporada en el cuerpo de teoría sobre grupos) a partir de los cuales se deducen todas las propiedades del número "probabilidad" y se abre todo el campo de aplicaciones sobre la base de un soporte matemático formalmente estructurado.

1. Sea $P(E, A)$ una probabilidad asociada a un resultado A de un experimento E, se asume que $P(E, A) \geq 0$. Es decir que la probabilidad es un número no negativo.
2. La probabilidad de un resultado del que se tiene la seguridad absoluta de ocurrencia vale 1. $P(E, A) = 1$ nos dice que A es un resultado determinado por el experimento E. El conjunto de todos los resultados posibles suele llamarse el "espacio muestral" y notado con S. Se admite la seguridad absoluta de que uno de los eventos posibles del espacio muestral debe ocurrir, de modo que este axioma suele expresarse como $P(E, S) = 1$ o simplemente $P(S) = 1$.
3. Si dos resultados posibles A y B son mutuamente excluyentes, la probabilidad de que ocurra uno u otro de los resultados es la suma de las probabilidades, es decir que si A y B son mutuamente excluyentes $P(E, A \vee B) = P(E, A) + P(E, B)$.

Teoremas básicos

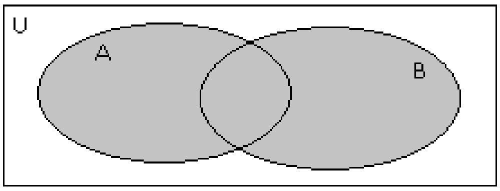
A partir de los postulados puede verse que el resultado A y la negación de su ocurrencia son mutuamente excluyentes (un resultado no puede obtenerse y no obtenerse simultáneamente). Por otra parte, la ocurrencia y no ocurrencia de A cubren todos los resultados posibles, por lo que se tiene la seguridad absoluta de que un evento A ocurre o no ocurre, sin que haya una tercera alternativa. Por lo tanto, si $\bar{A}$ representa la no ocurrencia de A, resulta

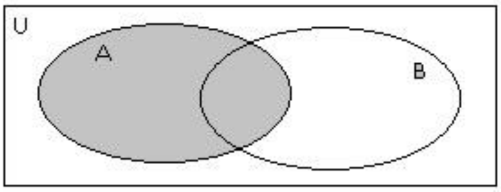
$$P(E, A \vee \bar{A}) = 1 = P(E, A) + P(E, \bar{A})$$

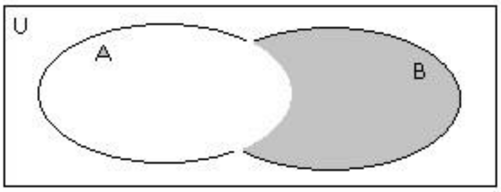
De aquí se deduce que

$$P(E,\bar{A}) = 1 - P(E,A)$$

En lo que sigue omitiremos la referencia al experimento E de la notación. Podemos deducir la expresión general de la probabilidad de la unión de un modo más formal considerando que esta operación entre dos eventos A y B puede escribirse en la forma $A \cup B = A \cup (B \cap \bar{A})$.







Como A y $B \cap \bar{A}$ son disjuntos, podemos afirmar que es válido $P(A \cup B) = P(A) + P(B \cap \bar{A})$. Por otra parte, el evento B puede escribirse como unión de dos eventos disjuntos en la forma $B = (B \cap A) \cup (B \cap \bar{A})$, de modo que podemos escribir $P(B) = P(B \cap A) + P(B \cap \bar{A})$. Si despejamos como diferencia de conjuntos el término $P(B \cap \bar{A}) = P(B) - P(A \cap B)$ y reemplazamos en la expresión de la unión, resulta

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

Vemos cómo, a partir de los axiomas, podemos establecer relaciones generales para el complemento y la unión de eventos.

En particular, si dos resultados son mutuamente excluyentes, la probabilidad de su ocurrencia simultánea es nula y la probabilidad de la ocurrencia de al menos uno de los resultados será la suma de las probabilidades de cada uno de ellos. Esto es válido para la probabilidad de la ocurrencia alternativa de muchos resultados excluyentes, que será la suma de las probabilidades individuales.

Probabilidades y cardinal de un conjunto

Vemos que, si se hace uso de las nociones de conjunto, puede accederse a un planteo más completo del problema del cálculo probabilístico sin introducirse en un tratamiento matemáticamente formal.

Desde este punto de vista se suele definir como "espacio muestral" S al conjunto de casos posibles resultantes de un experimento aleatorio. Al tirar una moneda, el espacio muestral es $S = \{Cara;\ Ceca\}$. Al tirar un dado es un número entero entre 1 y 6 de puntos sobre cada cara. Es claro que en todo juego de azar el espacio muestral describe el conjunto de casos posibles y el cardinal del espacio muestral es el número de casos posibles.

$$cp = \#S$$

Este espacio muestral se puede subdividir en un cierto número de subconjuntos o conjunto de partes. Cada uno de estos subconjuntos o partes se llama "evento" y se indica con letras mayúsculas como *A*, *B*, *C*, etc. *A* puede ser el evento "Cara" al tirar una moneda. Puede ser "el seis" al tirar un dado. O puede ser un número "par de puntos" (2, 4 o 6). De allí que podemos hablar de la probabilidad del evento A en relación con el experimento E en la medida que se especifique el experimento y defina el evento. Si se asume que A es el evento favorable, será $cf = \#A$. Esto nos permite redefinir la probabilidad empírica como

$$p = \frac{\#A}{\#S}$$

Más allá de ser más elegante como definición que la mención de “casos”, nos permite aplicar herramientas de la teoría de conjuntos al cálculo de probabilidades. Sin embargo, resulta difícil interpretar la noción de espacio muestral y de evento en términos de la definición de probabilidad como razón entre frecuencias.

En algunos casos es claro cuál es el espacio muestral fuera del marco de los juegos de azar, por ejemplo, cuando se define el experimento como la determinación del sexo de un niño antes de nacer. Pero en otros casos, como la distancia a la cual puede caer una piedra arrojada con la mano, es muy difícil definir de antemano el espacio muestral y se ve obligado a poner una cota arbitraria o bien admitir formalmente un infinito.

Puede verse también que los primeros ejemplos se referían a espacios muestrales discretos, con casos posibles aislados y bien identificados, mientras que el último ejemplo involucra un continuo de distancias posibles de caída.

En particular, si aplicamos estas nociones de conjunto al cálculo de probabilidades empíricas, podemos ver que, si n eventos son mutuamente excluyentes y simétricamente posibles, $P(A_i) = p_i = 1/\,n$ para todos los eventos. Si f de ellos son considerados exitosos o favorables, resulta $P(A) = f/n$, con lo cual se obtiene la definición clásica o empírica de probabilidad.

A partir de la noción de razón de frecuencias $p = f/n$, puede verse que no es posible tener un número negativo de casos favorables ni tampoco un número mayor que el de casos posibles. De allí que p es un número entre 0 y 1 $(0 \leq p \leq 1)$ siendo $p = 0$ la seguridad absoluta acerca de la imposibilidad de ocurrencia del resultado (ningún caso favorable) y $p = 1$ la seguridad absoluta de ocurrencia del resultado (todos los casos posibles son favorables). Estos dos límites representan extremos determinísticos. Todo el rango intermedio $(0 < p < 1)$ son situaciones probabilísticas. Notemos que desde la definición empírica (casos favorables sobre casos posibles), los extremos determinísticos son absolutamente seguros dado que decimos que no existe ningún caso favorable o que todos lo son, de modo que no hay posibilidad de falla en la predicción de éxito o fracaso. En la definición experimental o estadística (límite de frecuencia de éxitos al número de repeticiones), dado que se apoya en las repeticiones realizadas, que siempre se haya tenido éxito (o fracaso) no asegura que la próxima repetición también será exitosa. Esta observación tiene importancia en estadística cuando el límite teórico debe interpretarse en el marco del número efectivo de repeticiones de un experimento.

Otro resultado simple se obtiene intuitivamente de suponer que f_A representa la frecuencia de ocurrencias de *A*, f_B la frecuencia de ocurrencias de *B*, $f_{\text{AˆB}}$ la ocurrencia simultánea de *A* y de *B*, y $f_{\text{AˇB}}$ la ocurrencia de al menos uno de los resultados *A* o *B*. Por lo tanto f_A contiene los casos de ocurrencias de *A* así como los de ocurrencia de *A* y *B*. La frecuencia f_B contiene los casos de ocurrencia de *B* pero también los de *A* y *B*. Quiere decir que si se escribe a la frecuencia de ocurrencia de al menos uno de los eventos como $f_{\text{AˇB}} = f_A + f_B$, se estará sumando dos veces la ocurrencia simultánea de *A* y *B*, lo que no sería correcto. Por ello debe restarse una de las dos ocurrencias simultáneas para restablecer la igualdad. Resulta

$$f_{\text{AˇB}} = f_A + f_B - f_{\text{AˆB}}$$

Dividiendo por el número total de repeticiones *n* en ambos miembros queda

$$P(E, A\check{}B) = P(E, A) + P(E, B) – P(E, A\hat{}B)$$

Este planteo puede generalizarse a muchos resultados posibles, pero la expresión resulta más compleja.

Probabilidad condicional

La noción de probabilidad condicional es una de las más importantes dentro del cálculo de probabilidades porque encierra la relación entre la ocurrencia de eventos. Esta relación puede estar asociada con una vinculación causal, pero no necesariamente.

Se dice que un evento está condicionado por otro cuando la probabilidad de ocurrencia de tal evento se ve modificada por la ocurrencia o no del otro. Por ejemplo, imaginemos un juego que consiste en extraer tantas cartas de un mazo como indique el número resultante de tirar un dado. La probabilidad de obtener un as en las cartas dependerá del número resultante de la tirada del dado porque será mayor cuanto más grande sea ese número; la probabilidad de ocurrencia de accidentes depende de las condiciones de visibilidad y podemos seguir explorando ejemplos en los que la probabilidad de ocurrencia de un evento se ve modificada por la presencia o no de otro.

En términos generales, diremos que $P((E;A)/(E';B))$ representa la probabilidad condicional de que ocurra un evento A resultante de un experimento E si se tiene conocimiento acerca de la ocurrencia de un evento B resultante de un experimento E'. En forma abreviada se suele escribir $P(A/B)$ como "probabilidad de ocurrencia del evento A dado que se dispone de la información sobre la ocurrencia efectiva del evento B", o aún más abreviadamente "probabilidad de A dado B". Debe tenerse presente que lo que se evalúa es la probabilidad de ocurrencia de A, no se cuestiona la probabilidad de ocurrencia de B, sino que a B se lo da por conocido. También debe observarse que B no tiene que ser un evento cuya ocurrencia sea necesariamente previa a la de A –no se trata de una relación temporal- sino que tengamos el *conocimiento* de la ocurrencia de B *previamente* a la ejecución del experimento A. Por ejemplo, no se distingue la probabilidad de que alguien vaya a llegar tarde sabiendo que llueve, que vaya a llegar tarde sabiendo que llovió o que vaya a llegar tarde sabiendo que va a llover. La lluvia se asume como un dato y el problema se limita a la probabilidad de llegar tarde bajo condiciones de lluvia.

Desde el punto de vista del cálculo, la probabilidad condicional se relaciona de manera muy sencilla con las probabilidades usuales. Si f_{AB} representa la frecuencia de ocurrencias de A y B simultáneamente y f_B las de ocurrencia de B, es claro que la probabilidad condicional de A teniendo como dato la ocurrencia de B presupone la ocurrencia de ambos eventos, A y B. Pero el cálculo de probabilidad como cociente de frecuencias debe limitarse a la ocurrencia de B porque es un dato que se admite conocido. Por lo tanto f_B es el subconjunto de casos posibles de un total de *n*, el subconjunto que contiene la ocurrencia de B. Dentro de este subconjunto interesa el subconjunto aún más restringido de casos en que también ocurre A y que denotamos por f_{AB}. El cociente f_{AB}/f_B tenderá a la probabilidad condicional de A sabiendo que ocurre B cuando *n* tienda a infinito. Pero si dividimos numerador y denominador por el número total de casos *n* resulta

$$\frac{f_{AB}}{f_B} = \frac{\frac{f_{AB}}{n}}{\frac{f_B}{n}} = \frac{P(A \cap B)}{P(B)} = P(A/B)$$

Admitiremos por lo tanto en adelante que

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

será a la vez una definición y una forma genérica de calcular la probabilidad condicional de un experimento.

Puede notarse que la noción de dependencia impone que $P(A/B) \neq P(A/\bar{B})$, donde $\bar{B}$ representa el complemento de B e indica que el evento B no ha ocurrido, es decir que la información relativa al evento B altera la probabilidad de ocurrencia de A. Se infiere que B contiene información sobre la posible ocurrencia de A y por lo tanto hay dependencia.

Independencia

De acuerdo con lo planteado más arriba, la dependencia se manifiesta en la alteración del valor de una probabilidad asociada a un experimento condicionado por la ocurrencia de otro evento. Por ejemplo, es claro que la probabilidad de que una persona seleccionada al azar sea alta será menor que si la selección se hace entre los miembros de un equipo de basket. En este caso el conocimiento de la pertenencia de una persona, que nos van a presentar, a un equipo de basket nos permite intuir que debería ser alta, es decir que nos resulta más probable que sea alta que si se tratase de una persona de quién sabemos que no juega al basket.

Si partimos de la expresión escrita más arriba como definición, se deduce que la probabilidad de la intersección entre eventos puede escribirse como un producto $P(A \cap B) = P(A/B)P(B)$, lo que se suele conocer como el "teorema de la multiplicación". Pero por conmutatividad de la intersección $P(A \cap B) = P(B \cap A)$, que por simetría en la escritura del teorema nos permite escribir el teorema de la multiplicación en la forma $P(B \cap A) = P(B/A)P(A)$. Por lo tanto

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

Esta expresión será usada más adelante. Por el momento observamos que resulta

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

De modo que si P(A/B) ≠ P(A) es equivalente a decir que A depende de B, ahora debemos afirmar que es el evento B dependiente del evento A. Por lo tanto, se concluye que los eventos A y B son mutuamente dependientes. De modo que no puede establecerse un orden temporal ni causal sino en relación con que la información acerca de la ocurrencia de un evento contiene información en relación con la ocurrencia del otro.

Por otra parte, si A es independiente de B debe ser $P(A/B) = P(A)$. De la expresión de arriba resulta que $P(A \cap B) = P(A/B)P(B) = P(A)P(B)$. Como a su vez vale la relación $P(A)P(B) = P(B/A)P(A) = P(B \cap A) = P(B)P(A)$, debe ser $P(B/A) = P(B)$, es decir que B es independiente de A y resultan mutuamente independientes.

En otras palabras, las relaciones de dependencia e independencia no reconocen relación de orden. Los eventos son mutuamente independientes o mutuamente dependientes. Por lo tanto, no puede establecerse una relación de causalidad ni secuencia temporal a partir de la verificación de una relación de dependencia probabilística.

Recordemos aquí que si los eventos son independientes puede escribirse también la relación $P(A \cap B) = P(A)P(B)$, mientras que si los eventos son mutuamente excluyentes es válida la expresión $P(AUB) = P(A) + P(B)$. Así la dependencia y la exclusión facilitan mucho el cálculo de probabilidades en diferentes contextos. La probabilidad de la intersección de varios eventos puede escribirse como un producto de probabilidades si los eventos son todos mutuamente independientes, cada una de las cuales puede calcularse por separado. Si por otra parte se trata de eventos mutuamente excluyentes, la probabilidad de la unión de tales eventos se puede calcular como la suma de probabilidades que pueden calcularse por separado. Además, si se sabe que los eventos son mutuamente excluyentes, la ocurrencia de uno hace imposible la ocurrencia simultánea del otro, por lo que $P(A/B) = 0$ si A y B son excluyentes. Por otra parte, si A y B son independientes, no podemos decir nada acerca de que sean o no excluyentes.

Teoremas de la probabilidad total y de Bayes

Entre un conjunto de teoremas que de una u otra manera facilitan el cálculo de probabilidades, los teoremas de la probabilidad total y de Bayes están entre los más usados y conocidos.

El teorema de la probabilidad total propone que si se da la situación en la cual no es conocida la probabilidad de un evento, pero sí sus probabilidades condicionales en función de algún otro conjunto de eventos en que pueda subdividirse el espacio muestral a modo de partición, la probabilidad del evento puede encontrarse como una combinación lineal de las probabilidades condicionales en función de las probabilidades de los eventos de la partición.

Se entiende como partición a un conjunto de subconjuntos que conforman un conjunto mayor tales que, si *B* es el conjunto y $B_1 \dots B_n$ los elementos de la partición, sean

$$B = B_1 \cup B_2 \cup \dots \cup B_i \dots \cup B_j \dots \cup B_n \;\; (Reconstruyen\ el\ espacio\ muestral)$$
$$Para\ todo\ i, B_i \neq \emptyset\ (Ningún\ subconjunto\ es\ vacío)$$
$$Para\ par\ todo\ i\ y\ j\ B_i \cap B_j = \emptyset\ (Son\ mutuamente\ disjuntos)$$

Sea A un evento y $B_1 \dots B_n$ una partición de un espacio muestral B tal que se conocen las probabilidades condicionales $P(A/B_1) \dots P(A/B_i) \dots P(A/B_n)$, como el evento puede reconstruirse por medio de la unión de los fragmentos correspondientes a cada evento de la partición

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_i) \cup \dots \cup (A \cap B_n)$$

siendo cada uno de los subconjuntos entre paréntesis disjuntos entre sí, la probabilidad de A puede escribirse como una sumatoria

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i)$$

Pero a su vez cada intersección puede escribirse

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

resulta que

$$P(A) = \sum_{i=1}^{n} P(A/B_i)P(B_i)$$

que es la expresión usual del teorema de la probabilidad total.

Por ejemplo, si de tres máquinas A, B y C se producen 5%, 3% y 2% de artículos fallados respectivamente con proporciones de fabricación de 30%, 20% y 50% del total de los artículos, la probabilidad de que un artículo al azar se obtenga de la máquina A valdrá 0,3, de la B valdrá 0,2 y será de 0,5 para la máquina C. De allí que la probabilidad de que un artículo al azar, proveniente de cualquiera de las tres máquinas suponiendo que las tres producciones se mezclan de manera indistinguible, esté fallado (evento F) valdrá, de acuerdo con el teorema,

$$P(F) = P(F/A)P(A) + P(F/B)P(B) + P(F/C)P(C) =$$
$$= 0.05 * 0.3 + 0.03 * 0.2 + 0.02 * 0.5 =$$
$$= 0.015 + 0.006 + 0.01 = 0.031 = 3.1\%$$

El teorema de Bayes parte de la relación que fue planteada más arriba

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

Si reemplazamos a *B* por B_i, uno de los elementos de la partición, será

$$P(A/B_i)P(B_i) = P(B_i/A)P(A)$$

de donde, aplicando el teorema de la probabilidad total a la determinación de *P(A)*

$$P(B_i/A) = \frac{P(A/B_i)P(B_i)}{P(A)} = \frac{P(A/B_i)P(B_i)}{\sum_{i=1}^{n} P(A/B_i)P(B_i)}$$

Por ejemplo, si en el planteo anterior se ha calculado $P(F)$ por el teorema de la probabilidad total, puede saberse cuánto vale la probabilidad de que un elemento fallado provenga de la máquina C. Así $P\left(\frac{C}{F}\right) = \frac{P(F/C)P(C)}{P(F)} = \frac{0.02*0.5}{0.031} = 0.323$

Aplicación de los teoremas a técnicas de conteo

La técnica de conteo de permutaciones se resume a la independencia entre selecciones sucesivas con respecto a los resultados anteriores, con la excepción que cada elemento sólo puede ser ubicado una única vez en cada posición. Si $1/n$ es la probabilidad de que un elemento sea ubicado en el primer lugar, $1/(n-1)$ de que uno de los restantes sea ubicado en el segundo lugar, y así sucesivamente, resulta que la probabilidad de que se obtenga un ordenamiento determinado vale $1/n!$.

Un planteo similar puede hacerse en términos de variaciones y combinaciones. En el caso de permutaciones con repetición, cada elemento tiene una probabilidad de $1/n$ de ser ubicado en *r* posiciones diferentes, de donde $1/n^r$ es la probabilidad de un ordenamiento en particular.

Si p es la probabilidad de éxito, será *1-p* la probabilidad de fracaso en un experimento. Si este experimento se repite *n* veces de manera independiente con respecto a los resultados previos, que se obtenga *n* éxitos tendrá una probabilidad asociada de valor p^n, mientras que la probabilidad de que se obtenga *n* fracasos valdrá $(1-p)^n$. Si se cuestiona el valor de la probabilidad de obtener al menos un fracaso en *n* repeticiones, será $1-p^n$, mientras que obtener al menos un éxito valdrá $1-(1-p)^n$.

Si un experimento se repite *n* veces con independencia entre repeticiones, la probabilidad de obtener *r* éxitos y *n-r* fracasos valdrá, en principio $p^r(1-p)^{n-r}$. Pero en este planteo falta considerar el número de formas diferentes en las cuales ubicar los *r* éxitos indistinguibles entre *n* repeticiones. El número de formas posibles de ubicarlos, equivalentes en términos del número de éxitos y fracasos está dado por el combinatorio $nCr = \binom{n}{r}$, es decir, de cuántas maneras diferentes se pueden seleccionar *r* lugares entre *n* posiciones donde ubicar los éxitos (nótese que aquí se seleccionan lugares y no elementos). De esta manera se obtiene

$$P(r\ \acute{e}xitos\ y\ n-r\ fracasos) = \binom{n}{r} p^r (1-p)^{n-r}$$

Esta expresión se conoce como “distribución binomial o de Bernoulli” y será estudiada en detalle más adelante. El nombre de “binomial” proviene del conocido Binomio de Newton

$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^r b^{n-r}$$

Puede escribirse $1 = p + (1-p)$ y considerar $a = p$ y $b = 1-p$. Por lo tanto, la fórmula binomial se reduce a la expresión de arriba y puede verse que la suma sobre todos los casos posibles (entre 0 y n) da 1, como era de esperarse de considerar todos los casos posibles.

Probabilidades y Estadística: Ejercicios Unidad Nº 2

Cálculo Elemental de Probabilidades y aplicaciones de teoremas básicos

1. Supóngase que el conjunto universal U está dado por U = $\{x \ / \ 0 \leq x \leq 2\}$ y que los conjuntos A y B son A = $\{x \ / \ 0{,}25 \leq x \leq 1\}$ y B = $\{x \ / \ 0{,}5 \leq x \leq 1{,}5\}$. Describa los conjuntos siguientes: a) $A \cup B$ b) $A \cap B$ c) $AU\overline{B}$ d) $A \cap \overline{B}$.

2. Hallar la probabilidad de que un número entre 1 y 30 sea múltiplo de tres, de que sea par, de que sea múltiplo de tres y par, de que sea múltiplo de tres o par, de que no sea múltiplo de tres, de que no sea par, de que no sea múltiplo de tres ni par y de que no sea ambas cosas simultáneamente.

3. Una compañía eléctrica ofrece una tasa subsidiada a cualquier familia cuyo consumo de electricidad sea menor que 240 kWh durante un mes. Señalemos como A el evento en que una familia seleccionada al azar, en cierta comunidad, no rebasa el consumo subsidiado durante enero y como B el evento análogo para el mes de julio (A y B se refieren a la misma familia). Supongamos que P(A) = 0.8, P(B) = 0.7 y P(AUB) = 0.9. Calcule $P(A \cap B)$ y la probabilidad de que el consumo subsidiado sea rebasado uno de los dos meses, pero no el otro.

4. Si se tiran dos dados, hallar la probabilidad de obtener un seis en ambos resultados. Hallar la probabilidad de que no salga seis en ninguno de los dados. Hallar la probabilidad de obtener al menos un seis en las dos tiradas.

5. Se desea hallar: a) la probabilidad de sacar por lo menos una vez el 6 al arrojar un dado cuatro veces; b) la probabilidad de sacar dos 6 por lo menos una vez al lanzar los dados veinticuatro veces.

6. Se tira seis veces un dado. Calcular la probabilidad de que salga por lo menos un 6 en la secuencia de tiradas. ¿Qué ocurrirá si se tiran 36 veces dos dados con la probabilidad de que se obtengan por lo menos una vez dos 6 simultáneamente?

7. Un número binario está compuesto sólo por los dígitos 0 y 1. Supóngase que un número binario está formado por n dígitos, que la probabilidad de que aparezca un dígito incorrecto vale p y que los errores en dígitos diferentes son independientes entre sí. ¿Cuánto vale la probabilidad de formar un número incorrecto?

8. Una determinada travesía puede hacerse con aviones bimotores o con cuatrimotores. Los bimotores pueden volar con un solo motor y los cuatrimotores por lo menos con dos. La probabilidad de que un motor falle durante la travesía vale p. ¿Cuáles aviones son más seguros?

9. La ruta utilizada por cierto automovilista para ir al trabajo tiene dos cruces con semáforo. La probabilidad de que pare en el primer semáforo es de 0.4, la probabilidad análoga para el segundo es 0.5 y la probabilidad de que se detenga por lo menos en uno de los dos semáforos vale 0.6. ¿Cuál es la probabilidad de que se detenga en ambos semáforos?, ¿cuál es la probabilidad de que se detenga en el primero pero no en el segundo semáforo?, ¿cuál es la probabilidad de que se detenga exactamente en uno de ellos?

10. Una caja contiene cuatro tubos malos y cinco buenos. Se sacan dos tubos sucesivamente, se prueba el primero de ellos y comprueba que es bueno, calcular la probabilidad de que el otro también sea bueno. Si el primer tubo extraído fue malo, hallar la probabilidad de que el segundo tubo siga siendo bueno.

11. ¿En qué casos se puede escribir que P(AUB) = P(A) + P(B) y cuándo es correcto que $P(A \cap B) = P(A)\,P(B)$?

12. Considerando que los eventos independientes A, B y C poseen probabilidades Pa, Pb y Pc conocidas, interpretar las siguientes expresiones relativas al cálculo de probabilidades:

a) Pa (1 - Pb)
b) Pa Pb Pc
c) 1 - Pa Pb
d) Pa (1 - Pb) (1 - Pc)
e) 1 - Pa (1 - Pc)
f) Probar que $P(A \cup B \cup C)$ = Pa + Pb + Pc - Pa Pb - Pa Pc - Pb Pc + Pa Pb Pc.

13. En un tribunal de tres personas A, B y C, las probabilidades respectivas de dar un fallo justo son Pa = 0,8, Pb = 0,9 y Pc = 0,9. Se desea hallar la probabilidad de que el fallo del tribunal sea justo sabiendo que el mismo se toma por mayoría de votos.

14. Dos tubos defectuosos se confunden con dos buenos. Los tubos se prueban, uno por uno, hasta encontrar los dos defectuosos.

a) ¿Cuánto vale la probabilidad de encontrar los dos defectuosos en la segunda prueba?
b) ¿Cuánto vale la probabilidad de encontrar el último tubo defectuoso en la tercera prueba?
c) ¿Cuánto vale la probabilidad de encontrar el último tubo defectuoso en la cuarta prueba?
d) Sume las probabilidades obtenidas en los puntos anteriores y discuta el resultado.

15. Una urna numerada “1” contiene 10 esferas blancas y 20 esferas rojas. Otra urna “2” contiene 5 esferas blancas y 15 esferas rojas. Se escoge una esfera al azar de la urna “1” y se la pone en la urna “2”, entonces se toma una esfera al azar de la urna “2”. ¿Cuánto vale la probabilidad de que esta esfera, extraída de la urna “2” sea blanca? Hallar la probabilidad de que si la bolilla extraída de la segunda urna es blanca, lo haya sido también la que se extrajo de la urna 1.

16. Una máquina de tipo A produce 2% de artículos defectuosos y otra B un 5% de artículos fallados. Se dispone de 40 máquinas de tipo A y de 10 del tipo B, las cuales funcionan a la misma velocidad de producción. a) Hallar la probabilidad de que un artículo, proveniente de cualquiera de las máquinas, sea defectuoso. b) Si se toma un artículo al azar del total producido y encuentra que es defectuoso, calcular la probabilidad de que provenga de una máquina del tipo A.

17. Se tienen dos urnas numeradas 1 y 2, cada una con dos cajones. La urna 1 tiene una moneda de oro en un cajón y una de plata en el otro, mientras que la urna 2 tiene una moneda de oro en cada uno de los cajones. Se escoge una urna al azar y de esta un cajón al azar. La moneda encontrada en ese cajón resulta ser de oro. Calcular la probabilidad de que la moneda provenga de la urna 2.

18. Un sistema de bombeo consta de dos bombas idénticas 1 y 2. Si falla una bomba el sistema seguirá funcionando pero, debido al esfuerzo, ahora es más probable que la otra bomba falle que en la situación original. Es decir, r = P(falla 2 / falla 1) > P(falla 2) = q. Si en un año de operación hay 7% de fallas en una bomba y 1% de fallas en las dos, ¿cuánto vale la probabilidad individual de falla de cada bomba?

19. Una empresa que fabrica cámaras de video produce un modelo básico y un modelo de lujo. El año pasado 40% de las cámaras vendidas han sido del modelo básico. De los compradores del modelo básico 30% compran una garantía ampliada, mientras que el 50% de los compradores del modelo de lujo también lo hacen así. Si sabemos que un comprador seleccionado al azar tiene garantía ampliada, ¿qué tan probable es que tenga un modelo básico?

Resolución de los ejercicios de la Unidad 2

1. Supóngase que el conjunto universal U está dado por U = {x / 0 ≤ x ≤ 2} y que los conjuntos A y B son A = {x / 0,25 ≤ x ≤ 1} y B = {x / 0,5 ≤ x ≤ 1,5}. Describa los conjuntos siguientes: a) $A \cup B$ b) $A \cap B$ c) $A \cup \bar{B}$ d) $A \cap \bar{B}$.

(Se deja para actividad optativa como revisión de la noción de conjunto)

2. Hallar la probabilidad de que un número entre 1 y 30 sea múltiplo de tres, de que sea par, de que sea múltiplo de tres y par, de que sea múltiplo de tres o par, de que no sea múltiplo de tres, de que no sea par, de que no sea múltiplo de tres ni par y de que no sea ambas cosas simultáneamente.

Separamos los números entre el 1 y el 30 en los que son pares pero no múltiplos de 3 (2, 4, 8, 10, 14, 16, 20, 22, 26, 28), los que son múltiplos de 3 pero no pares (3, 9, 15, 21, 27), los que son pares y múltiplos de 3 (6, 12, 18, 24, 30), y los que no son pares ni múltiplos de 3 (1, 5, 7, 11, 13, 17, 19, 23, 25, 29). Llamemos a los pares A={2,4,6,8,10,12,14,16,18,20,22,24,26,28,30} y llamemos ahora al conjunto formado por los múltiplos de 3 B={3,6,9,12,15,18,21,24,27,30}. Basta contar para obtener que la P(A)=1/2, que P(B)=1/3. Si observamos en la primera clasificación que hay seis elementos en la intersección, es decir en los que son pares y múltiplos de tres, escribimos P(A∩B)=1/6.

Al pedir que sea múltiplo de tres o par, se refiere a la unión de los conjuntos indicados arriba. Basta contarlos para ver que son veinte elementos, por lo tanto, P(A∪B)=2/3. Pero también podemos aplicar la ley relativa a la unión de conjuntos

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{2} + \frac{1}{3} - \frac{1}{6} = \frac{4}{6} = \frac{2}{3}$$

De la misma manera podemos verificar la validez de la ley relativa al complemento para los que no son pares o no pertenecen al conjunto A, y para los que no son múltiplos de tres, o no pertenecen a B,

$$P(\bar{A}) = 1 - P(A) = 1 - \frac{1}{2} = \frac{1}{2} \quad y \quad P(\bar{B}) = 1 - P(B) = 1 - \frac{1}{3} = \frac{2}{3}$$

Que no sea múltiplo de 3 ni par refiere a los que quedan excluidos de ambos conjuntos. En la agrupación inicialmente propuesta hallamos diez elementos. Escribiríamos $P(\bar{A} \cap \bar{B}) = \frac{1}{3}$. Podemos recurrir a una de las leyes de De Morgan para trabajar un poco más sobre esta situación. Recordamos que

$$P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - \frac{4}{6} = \frac{2}{6} = \frac{1}{3}$$

Por otra parte, los que no son simultáneamente par y múltiplo de 3, refiere a no pertenecer a los elementos de la intersección. Vemos que hay 24 elementos que no cumplen a la vez las dos condiciones. Luego $P(\overline{A \cap B}) = \frac{5}{6}$. Como se trata de un complemento, podemos verificar que

$$P(\overline{A \cap B}) = 1 - P(A \cap B) = 1 - \frac{1}{6} = \frac{5}{6}$$

También podemos usar De Morgan

$$P(\overline{A \cap B}) = P(\bar{A} \cup \bar{B}) = P(\bar{A}) + P(\bar{B}) - P(\bar{A} \cap \bar{B}) = \frac{1}{2} + \frac{2}{3} - \frac{1}{3} = \frac{5}{6}$$

3. Una compañía eléctrica ofrece una tasa subsidiada a cualquier familia cuyo consumo de electricidad sea menor que 240kWh durante un mes. Señalemos como A el evento en que una familia seleccionada al azar, en cierta comunidad, no rebasa el consumo subsidiado durante enero y como B el evento análogo para el mes de julio (A y B se refieren a la misma familia). Supongamos que P(A)=0.8, P(B)=0.7 y P(AUB)=0.9. Calcule P(A∩B) y la probabilidad de que el consumo subsidiado sea rebasado uno de los dos meses, pero no el otro.

A partir de la probabilidad de la unión de conjuntos obtenemos

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0{,}8 + 0{,}7 - 0{,}9 = 0{,}6$$

Como la unión puede construirse por medio de las partes excluyentes o "partición"

$$A \cup B = (\bar{A} \cap B) \cup (A \cap \bar{B}) \cup (A \cap B)$$

Por ser excluyentes

$$P(A \cup B) = P(\bar{A} \cap B) + P(A \cap \bar{B}) + P(A \cap B)$$

Que el consumo sea rebasado en uno de los meses pero no el otro responde a la probabilidad de la unión

$$P(\bar{A} \cap B) + P(A \cap \bar{B}) = P(A \cup B) - P(A \cap B) = 0{,}9 - 0{,}6 = 0{,}3$$

4. Si se tiran dos dados, hallar la probabilidad de obtener un seis en ambos resultados. Hallar la probabilidad de que no salga seis en ninguno de los dados. Hallar la probabilidad de obtener al menos un seis en las dos tiradas.

Considerando el total de pares ordenados que puede obtenerse entre el (1;1) hasta el (6;6), vemos que hay 36 pares posibles, o bien considerando permutaciones con repetición de seis resultados diferentes. Entre los 36 pares sólo uno es el par (6;6), de modo que la probabilidad de obtener un seis en ambos resultados será 1/36.

La probabilidad de que no salga seis en uno de los dados es el complemento de que salga seis, por lo tanto 5/6. Y que no salga seis en ninguno será el cuadrado de este cociente, por lo tanto 25/36. Más formalmente, si A={sale 6 al tirar un dado} representa al evento individual

$$P(A) = \frac{1}{6}$$

Luego

$$P(\bar{A}) = 1 - \frac{1}{6} = \frac{5}{6}$$

Que salga el par (6;6) será

$$P(6;6) = P(A)^2 = \left(\frac{1}{6}\right)^2 = \frac{1}{36}$$

Y

$$P(\bar{6};\bar{6}) = \left(1 - \frac{1}{6}\right)^2 = \left(\frac{5}{6}\right)^2 = \frac{25}{36}$$

La probabilidad de obtener al menos un seis refiere a que no salga un número distinto de seis en los dos, por lo tanto el complemento del evento cuya probabilidad se ha calculado, luego

$$P(al\ menos\ un\ seis\ en\ dos\ tiradas) = 1 - \left(1 - \frac{1}{6}\right)^2 = \frac{11}{36} = 0{,}30\hat{5}$$

Retendremos esta expresión para resolver los tres ejercicios que siguen. Puede notarse también que se puede plantear

$$P\{(6;*) \cup (*;6)\} = P(6;*) + P(*;6) - P(6;6) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}$$

5. Se desea hallar: a) la probabilidad de sacar por lo menos una vez el 6 al arrojar un dado cuatro veces; b) la probabilidad de sacar dos 6 por lo menos una vez al lanzar los dados veinticuatro veces.

Considerando que la expresión del ejercicio 4 correspondía a obtener al menos un seis en dos tiradas, el mismo razonamiento aplicado a cuatro tiradas nos lleva a

$$P(al\ menos\ un\ seis\ en\ cuatro\ tiradas) = 1 - \left(1 - \frac{1}{6}\right)^4 = 0{,}51775$$

Notemos que un "razonamiento lineal" nos lleva a pensar que si multiplicamos por dos el número de dados, y por lo tanto multiplicamos por seis el número de opciones, si multiplicamos por seis el número de tiradas, debería obtenerse la misma probabilidad de éxito que en el cálculo anterior. Éste fue el razonamiento "lineal" que hizo el Caballero De Meré para plantearle el problema a Pascal dado que, cuando realizaba el experimento con un dado tirado cuatro veces y dos dados tirados veinticuatro veces, no obtenía la misma frecuencia de éxitos, como "debía ser". Pascal habría resuelto el problema y esto a su vez despertado en Pascal el interés por el cálculo de probabilidades. De acuerdo con nuestro planteo, para el caso de tirar dos dados veinticuatro veces debemos obtener

$$P(al\ menos\ dos\ seis\ en\ veinticuatro\ tiradas) = 1 - \left(1 - \frac{1}{36}\right)^{24} = 0{,}4914$$

Es claro que la solución al problema era altamente no lineal.

6. Se tira seis veces un dado. Calcular la probabilidad de que salga por lo menos un 6 en la secuencia de tiradas. ¿Qué ocurrirá si se tiran 36 veces dos dados con la probabilidad de que se obtengan por lo menos una vez dos 6 simultáneamente?

Sobre la base del ejercicio anterior

$$P(al\ menos\ un\ seis\ en\ seis\ tiradas) = 1 - \left(1 - \frac{1}{6}\right)^6 = 0{,}665$$

Y

$$P(al\ menos\ dos\ seis\ en\ treinta\ y\ seis\ tiradas) = 1 - \left(1 - \frac{1}{36}\right)^{36} = 0{,}637$$

7. Un número binario está compuesto sólo por los dígitos 0 y 1. Supóngase que un número binario está formado por n dígitos, que la probabilidad de que aparezca un dígito incorrecto vale p y que los errores en dígitos diferentes son independientes entre sí. ¿Cuánto vale la probabilidad de formar un número incorrecto?

Si seguimos adelante con el razonamiento anterior, para que haya una falla en la comunicación a través de una cadena de caracteres binarios, basta que al menos uno de los dígitos esté alterado. Si vale *p* la probabilidad de falla individual, será *1-p* la probabilidad de que la transmisión sea correcta, por lo tanto $(1-p)^n$ será la probabilidad de que todos los dígitos sean correctamente transmitidos y, por lo tanto, que la cadena sea correcta, mientras que $1-(1-p)^n$ será la probabilidad de que al menos uno de los dígitos sea incorrecto y en consecuencia toda la transmisión haya fallado.

8. Una determinada travesía puede hacerse con aviones bimotores o con cuatrimotores. Los bimotores pueden volar con un solo motor y los cuatrimotores por lo menos con dos. La probabilidad de que un motor falle durante la travesía vale p. ¿Cuáles aviones son más seguros?

Es ejercicio sintetiza un conjunto de elementos teóricos. Para empezar, si consideramos que un bimotor vuela con al menos uno de los motores, debemos evaluar la probabilidad de que funcionen los dos o de que al menos funcione uno. Es más sencillo calcular la probabilidad complementaria dado que sólo hay una única manera en que el bimotor no pueda volar, y es que fallen los dos. Sea p la probabilidad de falla individual de un motor. Como se supone que ambos motores funcionan en forma independiente, la probabilidad de que los dos fallen será p^2.

La misma evaluación para el cuatrimotor lleva a plantear la probabilidad de que fallen los cuatro motores, que será p^4, pero también pueden fallar tres motores y funcionar uno, y como hay cuatro maneras en que pueden fallar tres y funcionar uno (BMMM, MBMM, MMBM, MMMB), planteamos $4p^3(1-p)$ para expresar las cuatro maneras posibles en que fallan tres motores (p^3) y uno no falla $(1-p)$. Como es excluyente que fallen los cuatro con respecto a que fallen tres y uno no falle, la probabilidad de que el cuatrimotor no vuele será $p^4+4p^3(1-p)$.

Para decidir cuál avión es más seguro, basta comparar las expresiones obtenidas para la probabilidad de falla de cada uno para distintos valores de p en el rango de cero a uno. En forma de desigualdad, si

$$p^2 <=> p^4+4p^3(1-p)$$

Si $p=0$, ambos aviones son igualmente seguros porque los motores son perfectos. Si $p=1$ también vale la igualdad y los aviones son igualmente inseguros porque los motores no funcionan. Omitiendo los extremos, vemos que, como $p \neq 0$, podemos simplificar la relación de desigualdad y expresar el problema como

$$1 <=> p^2+4p(1-p)=p^2+4p-4p^2=4p-3p^2$$

Es decir, ver si

$$3p^2-4p+1 <=> 0$$

Si es menor que cero, los bimotores son más seguros, si es mayor que cero, los cuatrimotores son más seguros, y si es igual a cero, los dos son igualmente seguros. Planteemos la ecuación cuadrática y veamos si tiene solución en el intervalo de valores posibles para el número probabilidad. Al resolverla obtenemos $p=1/3$. Esto nos dice que si la probabilidad de falla vale 1/3, ambos aviones son igualmente seguros. Pero eso querría decir que uno de cada tres motores falla. Es razonable pensar que la probabilidad de falla es mucho menor que 0,3333. Como es la única raíz en el intervalo de interés dado que la otra raíz vale $p=1$, ya excluida en el razonamiento anterior, basta elegir un número menor que un tercio para la probabilidad y comparar el resultado con cero. Sea $p=0{,}1$ para verificar que 3*0,01-4*0,1+1=0,63>0. Esto nos dice, volviendo al planteo original, que la probabilidad de falla del cuatrimotor es menor que la del bimotor para valores razonables de probabilidad de falla de los motores en forma individual.

Este planteo puede hacerse también gráficamente dibujando las curvas de probabilidad de falla de cada avión en el intervalo de valores permitidos para el número probabilidad.

9. La ruta utilizada por cierto automovilista para ir al trabajo tiene dos cruces con semáforo. La probabilidad de que pare en el primer semáforo es de 0.4, la probabilidad análoga para el segundo es 0.5 y la probabilidad de que se detenga por lo menos en uno de los dos semáforos vale 0.6. ¿Cuál es la probabilidad de que se detenga en ambos semáforos?, ¿cuál es la probabilidad de que se detenga en el primero pero no en el segundo semáforo?, ¿cuál es la probabilidad de que se detenga exactamente en uno de ellos? Decidir si los semáforos funcionan independientemente uno del otro.

En este ejercicio se propone trabajar sobre los teoremas de la unión, complemento y la partición de conjuntos. La probabilidad de que pare en al menos uno de los dos semáforos responde a la unión de eventos. Si A corresponde a detenerse en el primer semáforo y B a parar en el segundo

$$0{,}6 = P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0{,}4 + 0{,}5 - P(A \cap B)$$

De aquí que $P(A \cap B) = 0{,}3$. Pero

$$A \cup B = (A \cap \bar{B}) \cup (A \cap B) \cup (\bar{A} \cap B)$$

Luego

$$P(A \cup B) = P(A \cap \bar{B}) + P(A \cap B) + P(\bar{A} \cap B)$$

De donde

$$P(A \cap \bar{B}) + P(\bar{A} \cap B) = P(A \cup B) - P(A \cap B) = 0{,}6 - 0{,}3 = 0{,}3$$

nos da la probabilidad de que se detenga en uno pero no en el otro.

Para calcular la probabilidad de que se detenga en el primero pero no en el segundo, lo planteamos como $P(A \cap \bar{B})$.

Observemos que

$$P(A) = P(A \cap B) + P(A \cap \bar{B})$$

Dado que

$$0{,}4 = 0{,}3 + P(A \cap \bar{B})$$

Luego

$$P(A \cap \bar{B}) = 0{,}1$$

Para decidir si funcionan independientemente evaluamos si la probabilidad de ocurrencia del evento A (detenerse en el primer semáforo) cambia cuando ocurre o no ocurre el evento B (se detiene o no se detiene en el segundo semáforo). Plantear entonces

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{0{,}3}{0{,}5} = 0{,}6$$

Por otra parte

$$P(A/\bar{B}) = \frac{P(A \cap \bar{B})}{P(\bar{B})} = \frac{0{,}1}{1 - 0{,}5} = 0{,}2$$

Al probar que las dos probabilidades son diferentes concluimos que el funcionamiento de los semáforos no es independiente.

10. Una caja contiene cuatro tubos malos y cinco buenos. Se sacan dos tubos sucesivamente, se prueba el primero de ellos y comprueba que es bueno, calcular la probabilidad de que el otro también sea bueno. Si el primer tubo extraído fue malo, hallar la probabilidad de que el segundo tubo siga siendo bueno.

Este ejercicio no es difícil desde el punto de vista numérico pero es engorroso si se pretende respetar la notación de condicionalidad. Podemos notar $P(A_1)$ como la probabilidad de extraer un tubo bueno en el primer intento, $P(A_2)$ para el segundo y las probabilidades condicionales $P(A_2/A_1)$ como la probabilidad condicional de extraer el segundo tubo bueno dado que el primer tubo fue bueno mientras que $P(A_2/\overline{A_1})$ representaría la probabilidad de que el segundo tubo haya sido bueno dado que el primero fue malo. El planteo resulta una aplicación inmediata del teorema de la multiplicación. Así

$$P(A_2/A_1) = \frac{4}{8} = 0{,}5 = 50\%$$

Mientras que

$$P(A_2/\overline{A_1}) = \frac{5}{8} = 0{,}625 = 62{,}5\%$$

11. ¿En qué casos se puede escribir que $P(A \cup B) = P(A) + P(B)$ y cuándo es correcto que $P(A \cap B) = P(A)P(B)$?

El planteo apunta a destacar, comparar y diferenciar dos situaciones muy ventajosas en términos de planteos de problemas en el ámbito del cálculo de probabilidades.

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

Esta relación es válida sólo si $P(A \cap B) = 0$, y esto ocurre si los eventos A y B son excluyentes, o, lo que es lo mismo, tienen intersección nula o no tienen elementos en común. Presenta el beneficio de establecer una "partición" del espacio muestral.

Por otra parte, dado que

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

Resulta

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

Esta relación es válida sólo si $P(A/B) = P(A) = P(A/\bar{B})$, lo que significa que los eventos A y B son independientes.

12. Considerando que los eventos independientes A, B y C poseen probabilidades Pa, Pb y Pc conocidas, interpretar las siguientes expresiones relativas al cálculo de probabilidades:

a) Pa (1 - Pb)
b) Pa Pb Pc
c) 1 - Pa Pb
d) Pa (1 - Pb) (1 - Pc)
e) 1 - Pa (1 - Pc)
f) Probar que P(A∪B∪C) = Pa + Pb + Pc - Pa Pb - Pa Pc - Pb Pc + Pa Pb Pc.

El objetivo del ejercicio es escribir las probabilidades expresadas numéricamente en términos de operaciones de conjuntos. Por ejemplo

a) $p_a(1 - p_b) = P(A \cap \bar{B})$
b) $p_a p_b p_c = P(A \cap B \cap C)$
c) $1 - p_a p_b = 1 - P(A \cap B) = P(\overline{A \cap B}) = P(\bar{A} \cup \bar{B})$
d) $p_a(1 - p_b)(1 - p_c) = P(A \cap \bar{B} \cap \bar{C})$
e) $1 - p_a(1 - p_c) = 1 - P(A \cap \bar{C}) = P\left(\overline{A \cap \bar{C}}\right) = P(\bar{A} \cup C)$
f) $P(A \cup B \cup C) = P\big(A \cup (B \cup C)\big) = P(A) + P(B \cup C) - P\big(A \cap (B \cup C)\big) = P(A) + P(B) + P(C) - P(B \cap C) - P\big((A \cap B) \cup (A \cap C)\big) = P(A) + P(B) + P(C) - P(B \cap C) - (P(A \cap B) + P(A \cap C) - P\big((A \cap B) \cap (A \cap C)\big) = P(A) + P(B) + P(C) - P(B \cap C) - P(A \cap B) - P(A \cap C) + P(A \cap B \cap C)$

13. En un tribunal de tres personas A, B y C, las probabilidades respectivas de dar un fallo justo son Pa = 0,8, Pb = 0,9 y Pc = 0,9. Se desea hallar la probabilidad de que el fallo del tribunal sea justo sabiendo que el mismo se toma por mayoría de votos.

Considerando que A, B y C representa el "evento fallo justo de los jueces A, B y C" respectivamente, con sus respectivos complemento, que al menos dos de los jueces sean justos se expresa

$$P(Fallo\ Justo) = P(A \cap B \cap C) + P(\bar{A} \cap B \cap B) + P(A \cap \bar{B} \cap C) + P(A \cap B \cap \bar{C})$$
$$= p_a p_b p_c + (1 - p_a) p_b p_c + p_a(1 - p_b) p_c + p_a p_b (1 - p_c) =$$

$$0{,}8.0{,}9.{,}9 + (1-0{,}8).0{,}9.0{,}9 + 0{,}8.(1-0{,}9).0{,}9 + 0{,}8.0{,}9.(1-0{,}9) = 0{,}954$$

14. Dos tubos defectuosos se confunden con dos buenos. Los tubos se prueban, uno por uno, hasta encontrar los dos defectuosos.

a) ¿Cuánto vale la probabilidad de encontrar los dos defectuosos en la segunda prueba?

b) ¿Cuánto vale la probabilidad de encontrar el último tubo defectuoso en la tercera prueba?

c) ¿Cuánto vale la probabilidad de encontrar el último tubo defectuoso en la cuarta prueba?

d) Sume las probabilidades obtenidas en los puntos anteriores y discuta el resultado.

Tomando la expresión del ejercicio 10, extraer los dos tubos defectuosos en la segunda extracción requiere que tanto el primero como el segundo tubo hayan sido defectuosos,

$$P(fin\ en\ 2°\ extracción) = P(\overline{A_2} \cap \overline{A_1}) = P(\overline{A_2}\ /\overline{A_1}) * P(\overline{A_1}) = \frac{2}{4} * \frac{1}{3} = \frac{2}{12} = \frac{1}{6}$$

$$P(fin\ en\ 3°\ extracción) = P(\overline{A_3} \cap A_2 \cap \overline{A_1}) + P(\overline{A_3} \cap \overline{A_2} \cap A_1) =$$
$$= P\left(\overline{A_3}/(A_2 \cap \overline{A_1})\right) * P(A_2 \cap \overline{A_1}) + P\left(\overline{A_3}/(\overline{A_2} \cap A_1)\right) * P(\overline{A_2} \cap A_1)$$
$$= P(\overline{A_3}/(A_2 \cap \overline{A_1})) * P(A_2/\overline{A_1}) * P(\overline{A_1}) + P\left(\overline{A_3}/(\overline{A_2} \cap A_1)\right) * P(\overline{A_2}/A_1)$$
$$* P(A_1) = \frac{1}{2} * \frac{2}{3} * \frac{2}{4} + \frac{1}{2} * \frac{2}{3} * \frac{2}{4} = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

La última la expresamos sólo numéricamente

$$P(fin\ en\ 4°\ extracción) = \frac{1}{1} * \frac{1}{2} * \frac{1}{3} * \frac{2}{4} + \frac{1}{1} * \frac{1}{2} * \frac{2}{3} * \frac{2}{4} + \frac{1}{1} * \frac{2}{2} * \frac{1}{3} * \frac{2}{4} = \frac{3}{6} = \frac{1}{2}$$

Que corresponde a las secuencias, leídas de derecha a izquierda

$$MBBM --- MBMB --- MMBB$$

Podemos notar que la suma de todas las opciones vale 1, como era de esperarse, pero lo más importante es ver cómo el análisis de la condicionalidad es complejo desde el planteo mismo, y sin embargo es uno de los problemas más relevantes en el cálculo de probabilidades y más aún en estadística, porque trata el vínculo entre eventos. En un primer abordaje trabajaremos sobre la hipótesis de independencia.

15. Una urna numerada “1” contiene 10 esferas blancas y 20 esferas rojas. Otra urna “2” contiene 5 esferas blancas y 15 esferas rojas. Se escoge una esfera al azar de la urna “1” y se la pone en la urna “2”, entonces se toma una esfera al azar de la urna “2”. ¿Cuánto vale la probabilidad de que esta esfera, extraída de la urna “2” sea blanca? Hallar la probabilidad de que si la bolilla extraída de la segunda urna es blanca, lo haya sido también la que se extrajo de la urna 1.

Este ejercicio propone una aplicación inmediata del teorema de la probabilidad total. Si escribimos que el evento A representa extraer una bolilla blanca y que B representa una bolilla roja, mientras que 1 y 2 refieren a las urnas 1 y 2 respectivamente, expresamos

$$P(A_2) = P(A_2 \cap A_1) + P(A_2 \cap B_1) = P(A_2/A_1) * P(A_1) + P(A_2/B_1) * P(B_1)$$
$$= \frac{6}{21} * \frac{10}{30} + \frac{5}{21} * \frac{20}{30} = 0{,}254$$

16. Una máquina de tipo A produce 2% de artículos defectuosos y otra B un 5% de artículos fallados. Se dispone de 40 máquinas de tipo A y de 10 del tipo B, las cuales funcionan a la

misma velocidad de producción. a) Hallar la probabilidad de que un artículo, proveniente de cualquiera de las máquinas, sea defectuoso. b) Si se toma un artículo al azar del total producido y encuentra que es defectuoso, calcular la probabilidad de que provenga de una máquina del tipo A.

Partimos del teorema de la probabilidad total para hallar la probabilidad de obtener un artículo defectuoso de cualquiera de las máquinas. Podemos escribirlo como en el ejercicio anterior o de otra manera para utilizar diferentes formas de planteo. Si referimos con "D" el evento "artículo defectuoso", y tomamos la partición como máquinas de tipo A (S_1) y máquinas de tipo B (S_2), $n=2$ tipos elementos en la partición, resulta

$$P(D) = \sum_{i=1}^{n} P(D/S_i)P(S_i) = P(D/A)P(A) + P(D/B)P(B) = 0{,}02 * \frac{40}{50} + 0{,}05 * \frac{10}{50}$$

Con lo cual P(D)=0,026=2,6%, y con esto respondemos al punto "a".
Para responder al punto "b" recurrimos al teorema de Bayes.

$$P(A/D) = \frac{P(D/A)P(A)}{P(D)} = \frac{0{,}02 * 40/50}{0{,}026} = 0{,}651 = 65{,}1\%$$

Al no haber más opciones, resulta más probable que le artículo defectuoso provenga de una máquina tipo A aunque producen menos defectuosos en proporción que las máquinas tipo B.

17. Se tienen dos urnas numeradas 1 y 2, cada una con dos cajones. La urna 1 tiene una moneda de oro en un cajón y una de plata en el otro, mientras que la urna 2 tiene una moneda de oro en cada uno de los cajones. Se escoge una urna al azar y de esta un cajón al azar. La moneda encontrada en ese cajón resulta ser de oro. Calcular la probabilidad de que la moneda provenga de la urna 2.

El ejercicio 17 es similar al 16 pero aún más simple dado que puede resolverse con un conteo. Usaremos sin embargo el teorema de Bayes. La partición está establecida por las urnas. Como la urna uno (U1) contiene una moneda de oro y una de plata en sendos cajones, la probabilidad de extraer la moneda de oro de la urna uno vale 0,5. Como la urna dos (U2) contiene dos monedas de oro, una en cada cajón, la probabilidad de extraer una moneda de oro de la urna dos vale 1. Como las dos urnas son equivalentes, sus respectivas probabilidades de elección valen 0,5. Luego

$$P(Oro) = P(Oro/U1)P(U1) + P(Oro/U2) * P(U2) = 0{,}5 * 0{,}5 + 1 * 0{,}5 = 0{,}75$$

Si se ha extraído una moneda de oro y quiere saber cuánto vale la probabilidad de que haya provenido de la urna dos calculamos

$$P(U2/Oro) = \frac{P(Oro/U2)P(U2)}{P(Oro)} = \frac{1 * 0{,}5}{0{,}75} = 0,\hat{6} = 66,\hat{6}\%$$

18. Un sistema de bombeo consta de dos bombas idénticas 1 y 2. Si falla una bomba el sistema seguirá funcionando pero, debido al esfuerzo, ahora es más probable que la otra bomba falle que en la situación original. Es decir, r = P(falla 2 / falla 1) > P(falla 2) = q. Si en un año de operación hay 7% de fallas en una bomba y 1% de fallas en las dos, ¿cuánto vale la probabilidad individual de falla de cada bomba?

Si llamamos $q = P(falla\ 2)$, como son idénticas, debe ser $q = P(falla\ 1)$. Por otra parte y por la identidad, $P(falla\ 1 \wedge \sim falla 2) = P(\sim falla\ 1 \wedge falla\ 2)$ (la probabilidad de que falle la bomba 1 y no falle la bomba 2 es igual a la probabilidad de que no falle la bomba 1 y falle la bomba 2). Sabemos que la suma de estas probabilidades (que falle una y la otra no, sin especificar cuál, por lo cual se refiere a las dos alternativas) vale 0,07 (7%). También se sabe que $0{,}01 = P(falle\ 1\ \hat{}\ falle\ 2)$. Estableciendo una partición a partir de la unión de eventos, tenemos que la probabilidad de la unión vale

$$P(falla\ 1\hat{}{\sim}falla\ 2) + P({\sim}falla\ 1\hat{}falla\ 2) + P(falla\ 1\hat{}falla\ 2) = 0{,}07 + 0{,}01$$

Como $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0{,}08 = q + q - 0{,}01$, de donde resulta $q = 0{,}045$. De este modo podemos calcular

$$r = P(falla\ 2\ /\ falla\ 1) = \frac{P(falla\ 2\ \hat{}\ falla\ 1)}{P(falla\ 1)} = \frac{0{,}01}{0{,}045} = 0{,}222\ ...$$

Esta probabilidad es claramente diferente de la probabilidad de que falle la bomba 2 en forma individual. Inclusive puede calcularse

$$P(falla\ 2\ /\ {\sim}falla\ 1) = \frac{P(falla\ 2\ \hat{}{\sim}\ falla\ 1)}{P({\sim}falla\ 1)} = \frac{0{,}035}{0{,}965} = 0{,}036$$

No sólo la falla de una bomba influye fuertemente sobre el riesgo de falla de la otra sino que, al saber que una no falla, la probabilidad de falla de la otra es apenas superior a la falla individual si opera sola.

19. Una empresa que fabrica cámaras de video produce un modelo básico y un modelo de lujo. El año pasado 40% de las cámaras vendidas han sido del modelo básico. De los compradores del modelo básico 30% compran una garantía ampliada, mientras que el 50% de los compradores del modelo de lujo también lo hacen así. Si sabemos que un comprador seleccionado al azar tiene garantía ampliada, ¿qué tan probable es que tenga un modelo básico?

El ejercicio es una aplicación del teorema de Bayes. Calculamos primero la probabilidad de que el comprador haya pedido una garantía ampliada

$$P(GA) = P(GA/MB)P(MB) + P(GA/ML)P(ML) = 0{,}3 * 0{,}4 + 0{,}5 * 0{,}6 = 0{,}42$$

En función del teorema de Bayes

$$P(MB/GA) = \frac{P(GA/MB)P(MB)}{P(GA)} = \frac{0{,}3 * 0{,}4}{0{,}42} = 0{,}286$$

Unidad 3: Variables Aleatorias y Distribuciones

Noción de variable aleatoria

La idea de usar variables aleatorias responde al desarrollo de un formalismo en el cálculo de probabilidades. Con este formalismo se pretende hacer uso de las herramientas matemáticas con el fin de sistematizar el aporte de un criterio para la toma de decisión bajo condiciones de incertidumbre.

Una variable aleatoria no es más que un número que identifica un evento. En tal sentido puede usarse cualquier número que identifique al evento de manera biunívoca. Por ejemplo, en lugar de "cara" o "ceca" puede identificarse los eventos posibles de tirar una moneda con 0 y 1, o bien con 1 y -1, o bien con los irracionales e y π. Los resultados posibles de tirar un dado se expresan con los números del 1 al 6, si bien en realidad son puntos en un cuadrado que también podrían ser colores sobre las distintas caras del dado. El resultado de una medición de distancia es una variable aleatoria como expresión del resultado mismo.

Vemos que las variables aleatorias pueden definirse en forma extensiva, indicando cada uno de los resultados posibles y el número asociado, o en forma intensiva mediante alguna ley de asignación, como el resultado de una medición.

En términos más formales, puede definirse una variable aleatoria como una función biyectiva que establece una relación biunívoca entre un evento y un número. Se denota una variable aleatoria mediante letras mayúsculas X, Y, Z, y sus resultados numéricos mediante letras minúsculas, como x, y, z. Así, la expresión $P(X = x_a) = P_{(A)} = p_a$ expresa que el evento *A*, con probabilidad p_a, también puede notarse por medio de $P(X = x_a)$ o "probabilidad de que la variable aleatoria X adquiera el valor particular x_a". En otras palabras, *A* es un evento de un espacio muestral *S*, *X* es una variable aleatoria que asigna un número a cada evento de acuerdo con algún criterio; finalmente x_a es el valor numérico de la variable aleatoria *X* correspondiente al evento *A*.

Proponemos dos grandes tipos de variables aleatorias: discretas y continuas. Ejemplos de variable aleatoria discreta (VAD) son las caras de una moneda o del dado, el número de personas que puede haber en un vehículo, el número de sillas en un aula, el número de estrellas en el cielo. Como se puede notar, se trata de números enteros o de eventos de un modo u otro "contables", aun cuando puedan ser infinitos. En términos más formales, si a un experimento aleatorio se puede asignar una variable aleatoria como función biyectiva con el conjunto de los naturales, tal variable aleatoria es discreta. Pero hay otros experimentos como la medida de una longitud, de un tiempo, de una masa o de una velocidad, que no admiten una biyección con los naturales porque responden a un rango continuo de números asignables. No nos referimos a la limitación práctica de medir en milímetros por problemas de resolución sino a la ausencia de límite teórico para medir una magnitud tal como una longitud.

Distribución puntual

Se entiende como distribución puntual de probabilidad al conocimiento de la probabilidad asociada a cada uno de los resultados posibles de un experimento, es decir, a la probabilidad asociada a cada uno de los valores numéricos que puede adoptar una variable aleatoria. En modo más formal, a conocer

$$P(X = x_i) = p_i \ \forall i$$

donde $p_i \geq 0 \ \forall i$ y $\sum_{i=1}^{n} p_i = 1$ sumado sobre todos los valore posibles de *i*. Es claro que si se dispone del valor numérico de probabilidad asociado a cualquier resultado posible, puede calcularse cualquier probabilidad en cualquier situación que se desee.

Es claro también que ésta definición tiene sentido si puede asignarse un valor de probabilidad a cada evento posible, y veremos que esto sólo es válido si la variable aleatoria es discreta. Antes de analizar las limitaciones y cómo se pueden superar, definamos la distribución acumulada.

Suele definirse como distribución puntual acumulada a la expresión aditiva

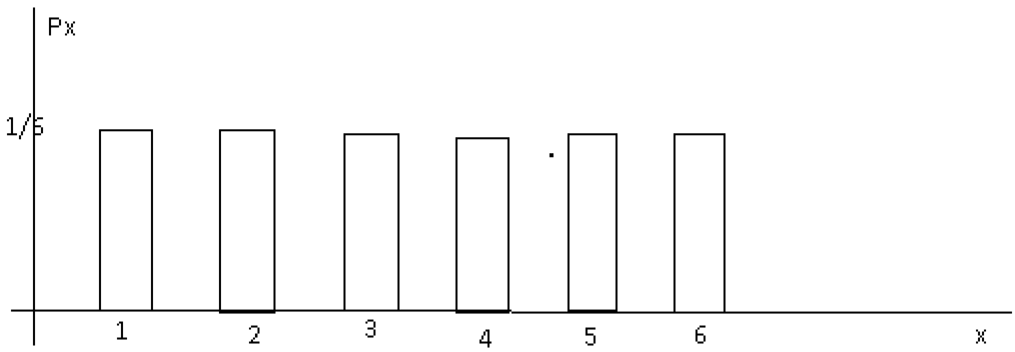
$$P(X \leq x_i) = \sum_{i=1}^{n} p_i \ \forall i$$

Es claro que este valor se obtiene como sumatoria de todas las probabilidades asociadas a cada resultado posible expresado en términos de variables aleatorias.

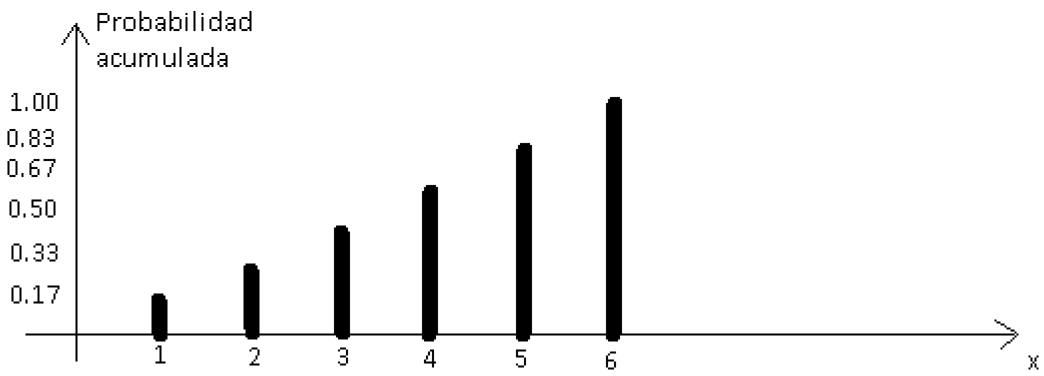
Al tirar un dado, si se define como variable aleatoria el número de puntos que quedan visibles en la cara superior del dado, se obtiene la siguiente distribución puntual y acumulada

X	$P(X = x_i)$	$P(X \leq x_i)$
1	1/6	1/6
2	1/6	2/6
3	1/6	3/6
4	1/6	4/6
5	1/6	5/6
6	1/6	6/6 = 1

Una expresión gráfica complementaria para visualizar el comportamiento de la distribución puntual puede ser



Un gráfico como el anterior sólo expresa que cada uno de los valores posibles de la variable aleatoria tiene la probabilidad informada en la tabla de valor 1/6. Para representar gráficamente la distribución acumulada podemos usar variantes gráficas o la misma estructura de representación, como en el gráfico que sigue.



En todos los casos hemos representado la probabilidad asociada a cada uno de los resultados posibles limitándonos al valor estrictamente indicado en el eje de abscisas. El resto de los puntos en el eje real son ignorados. Es claro que la distribución acumulada no contiene más información que la distribución puntual.

Al tirar dos monedas y contabilizar el número de caras, la distribución de probabilidad corresponde a la siguiente tabla

X	$P(X = x) = p_x$	$P(X \leq x) = p_{acx}$
0	0,25	0,25
1	0,5	0,75
2	0,25	1,00

Si tiramos dos dados y sumamos los resultados de las caras, vemos que pueden estar en el rango de 2 (dos unos) a 12 (dos seis). La distribución resulta de contabilizar el número de combinaciones de resultados que permite sumar cada uno de los valores posibles. Agregamos el "1", el "13" y el "14" para decir explícitamente que son resultados imposibles porque no hay ninguna combinación de formas en que puedan caer los dados que dé esa suma. Sabemos que hay 36 casos posibles pero sólo uno de ellos dará por resultado la suma $X=2$. Si sale "1" el primero y "2" el segundo o bien "2" el primero y "1" el segundo, la suma dará $X=3$, de modo que hay dos casos favorables entre 36 posibles para obtener como suma $X=3$. Si seguimos analizando los posibles resultados veremos que hay tres modos diferentes de obtener $X=4$ con un máximo de seis casos favorables para obtener $X=7$, el máximo de probabilidad de la distribución. La tabla resulta finalmente

X	$P(X = x) = p_x$	$P(X \leq x) = p_{acx}$
1	0	0
2	1/36	1/36
3	2/36	3/36
4	3/36	6/36
5	4/36	10/36
6	5/36	15/36
7	6/36	21/36
8	5/36	26/36
9	4/36	30/36
10	3/36	33/36
11	2/36	35/36
12	1/36	1=36/36
13	0	0
14	0	0

Se deja como ejercicio construir los gráficos de las distribuciones de probabilidad de los dos experimentos previos. Notemos que a la probabilidad acumulada para los valores numéricos 13 y 14 asignamos un valor de "cero" porque no forman parte del rango permitido de la variable aleatoria propuesta.

Distribución binomial o de Bernoulli

Una distribución discreta de uso muy habitual se conoce como "distribución binomial" o de "Bernoulli". Esta distribución fue propuesta un poco más arriba en términos de una forma práctica de calcular la probabilidad de obtener cierto número de éxitos y el resto de fracasos en n repeticiones independientes de un experimento con dos alternativas posibles.

Si llamamos p a la probabilidad de éxito de un experimento que se repite n veces en forma independiente. Si se obtiene r éxitos, calculamos esta probabilidad en la forma

$$p.p.p \dots p = p^r$$

Para expresar la probabilidad de que ocurran $n-r$ fracasos calculamos

$$(1-p)^{n-r}$$

El producto de las dos probabilidades expresa que ocurren r éxitos y $n-r$ fracasos. El número combinatorio contabiliza el número de maneras en que se puede seleccionar r ubicaciones para los éxitos entre las n repeticiones. La selección de *r* objetos (lugares para los éxitos) entre las *n* objetos (lugares para las repeticiones) distinguibles sin importar el orden está dada por el combinatorio $\binom{n}{r}$, de modo que el producto entre las formas diferentes de ubicar los éxitos y la probabilidad asociada nos da la distribución de probabilidad binomial o de Bernoulli. Esta distribución no sólo se expresa a través de una tabla y un gráfico sino a través de una expresión algebraica

$$P(r\ \acute{e}xitos\ y\ n-r\ fracasos)=\binom{n}{r}p^r(1-p)^{n-r}$$

Como ejemplo de esta distribución podemos considerar el cálculo de la probabilidad de que al tirar diez veces un dado se obtengan tres seis. Para ello, como las tiradas son independientes y no se altera las propiedades del dado en cada tirada, es aplicable una distribución binomial con probabilidad de éxito 1/6 y de fracaso 5/6. Por lo tanto, si X es el número de seis en las diez tiradas,

$$P_{(X=3)}=\binom{10}{3}\left(\frac{1}{6}\right)^3\left(1-\frac{1}{6}\right)^{10-3}=\frac{10!}{3!\,(10-3)!}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^7=0{,}23257$$

Se recomienda calcular el resto de las probabilidades asociadas a este caso particular, expresar la distribución puntual y acumulada en forma de una tabla, en formas de representación gráfica.

Función de distribución

El problema de las variables aleatorias continuas es que no admiten un valor puntual de probabilidad para ningún caso posible de los valores permitidos de una variable aleatoria. Desde el punto de vista de la definición “empírica” que usamos en la teoría de juegos, tenemos infinitos no numerables casos posibles y no podemos “dividir por infinito”. Desde el punto de vista de la noción de límite, sería como arrojar una piedra y preguntarnos por la probabilidad de que caiga exactamente a πm (pi metros) de distancia, con todos los valores decimales del número irracional π. Vemos que en ambos casos no tiene sentido la definición aplicada en la teoría de juegos ni la experimental.

Para resolver este problema podemos extender en cambio la definición de distribución acumulada “hasta cada uno de los valores posibles” de la variable aleatoria en un sentido más formal acotando superiormente el rango por “cada número real”. Sobre la base de la distribución acumulada se define la función de distribución $F(x)$

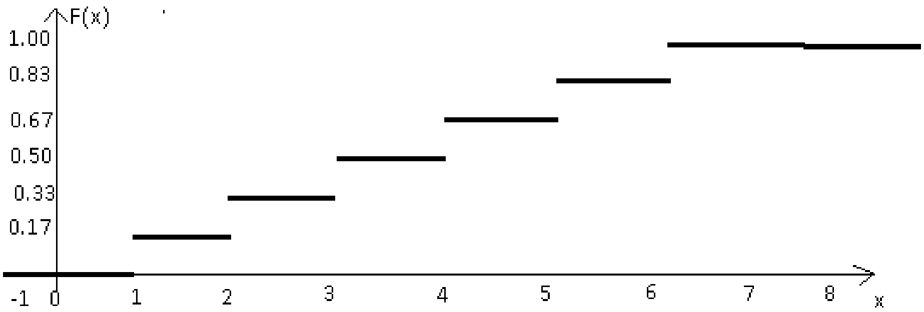
$$F(x)=P(X\leq x)\ \forall x\in\boldsymbol{R}$$

Sobre la base del ejemplo anterior, podemos construir la función de distribución correspondiente al experimento aleatorio “tirar un dado”.

Esta definición difiere de la anterior en que se trata de una función de números reales definida sobre todo el campo de los números reales. Por lo tanto nos habilita a usar herramientas del análisis de funciones aplicadas al cálculo de probabilidades, en particular se podrá aplicar a variables aleatorias continuas.

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{6} & 1 \leq x < 2 \\ \frac{2}{6} & 2 \leq x < 3 \\ \frac{3}{6} & 3 \leq x < 4 \\ \frac{4}{6} & 4 \leq x < 5 \\ \frac{5}{6} & 5 \leq x < 6 \\ \frac{6}{6} = 1 & 6 \leq x \end{cases}$$

En el siguiente gráfico se presenta la función de distribución correspondiente a “tirar un dado” de acuerdo con la descripción anterior



Vemos que numéricamente los valores de la función de distribución y de la distribución acumulada son iguales, pero mientras la distribución acumulada asigna valores puntuales en cada uno de los valores numéricos que pueden resultar de “tirar un dado”, en la función de distribución se representa con un gráfico escalonado con valores de *F(x)* para todos los números reales, no solamente los resultantes de tirar un dado.

Es claro que si se trata de una variable aleatoria discreta, será más útil disponer de la distribución puntual de probabilidad. Puede verse, sin embargo, que a partir de la distribución puntual puede construirse $F(x)$, que resultará una función escalón con saltos en las discontinuidades y constante donde la variable aleatoria no esté definida. Por otra parte, a partir de $F(x)$ puede construirse la distribución puntual. Si restamos al valor de probabilidad acumulada $F(x^*)$, donde x^* es una discontinuidad, el valor inmediatamente anterior, se obtiene el salto asociado a la discontinuidad, que coincide con el valor de probabilidad puntual asociado a x^*. Es decir, la diferencia entre el valor acumulado hasta x^* y el límite cuando *x* tiende a x^* por izquierda

$$P(X = x^*) = F(x^*) - \lim_{x \to x^{*-}} F(x)$$

En términos más generales, por ser *F* una distribución acumulada, será válida la relación $F(b) = F(a) + F(a < x \leq b)$. Por lo tanto

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

Se infiere que el conocimiento de *F(x)* determina completamente las probabilidades asociadas a un evento. Es claro también que *F(x)* es una función no decreciente. En los extremos del rango real $F(+\infty) = 1$ porque se ha acumulado toda la probabilidad posible; también $F(-\infty) = 0$ porque no se ha incorporado ningún evento dentro de los resultados posibles.

Es usual que se utilice el término “distribución de masa” al referirse a la distribución de probabilidad. Se hace uso en tal caso de una imagen física consistente en una superficie plana con la forma asociada a la distribución. En tal caso $P(X = x)$ se imagina como una barra, que puede ser de espesor infinitesimal, ubicada en el punto “x” y de masa “p” (probabilidad) de manera total que la masa total de todas las barras vale uno. En tal caso *F(x)* es la suma de todas las masas asociadas a los valores de variable aleatoria menores o iguales que x, resultando una función escalón con saltos en cada valor de x, y constante en cada intervalo entre un valor de x posible y el siguiente. Sobre la base de esta imagen se suele construir gráficos de funciones escalonadas para $F_{(x)}$ y gráficos de barra para las distribuciones puntuales, donde la abscisa es la variable aleatoria y la ordenada representa los valores de *p* o los saltos en $F_{(x)}$.

Extensión a variables aleatorias continuas

Hay eventos como el tiempo que tarda en desarrollarse un proceso y en general medidas de distancia, de masa, de temperatura que tienen asociado un rango continuo de valores posibles como variable aleatoria. En tales casos no es factible obtener una distribución puntual de probabilidades. En cambio tiene validez la función de distribución $F_{(x)}$ en tanto probabilidad acumulada. Si la variable aleatoria tiene un rango continuo, $F_{(x)}$ será una función continua no decreciente en tanto acumula probabilidad en cada punto del rango de la variable aleatoria.

En una variable aleatoria continua la noción empírica de probabilidad pierde sentido porque se trataría de un número finito de casos favorables entre infinitos posibles. Pero dentro del rango continuo puede analizarse la probabilidad asociada a un intervalo y, en el extremo puntual, resulta $P(X = x^*) = F_{(x^*)} - \lim_{x \to x^{*-}} F_{(x)}$, que coincide con $F_{(x^*)}$ si X es continua, por lo cual $P(X = x^*) = 0$. Esto se interpreta no como una imposibilidad de que resulte el valor x^* del experimento sino como la imposibilidad de conocer la probabilidad de ocurrencia de un valor puntual en un rango infinito de valores posibles.

Para ofrecer una idea más concreta y tratar de llegar “naturalmente” a la función de distribución continua y a la noción de función de densidad, tomemos como referencia el gráfico de la página anterior. Allí se indican los seis saltos de probabilidad correspondientes a los incrementos de probabilidad acumulada en un dado. Imaginemos que nuestro dado tiene doce caras con los números 0,5 a 6,0 en intervalos de 0,5, es decir que la nueva variable aleatoria sería

$$X = \{0.5; 1.0; 1.5; 2.0; 2.5; 3.0; 3.5; 4.0; 4.5; 5.0; 5.5; 6.0\}$$

Cada uno de estos doce valores posibles de la variable aleatoria tendría una probabilidad asociada de un doceavo, a diferencia de un sexto en un dado normal. La forma del grafico sería la misma, pero con doce saltos de un doceavo de probabilidad cada vez que avanzamos en 0,5 en la variable aleatoria.

Sigamos por este camino y construyamos un dado de sesenta caras. Nuestra nueva variable aleatoria será

$$X = \{0.1; 0.2; 0.3; \dots\dots\dots; 5.7; 5.8; 5.9;\ 6.0\}$$

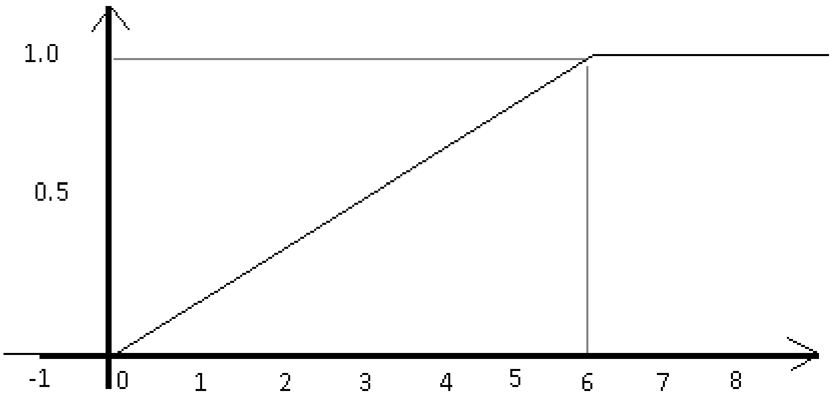
Tendremos sesenta saltos de 1/60 (un sesentavo) de probabilidad durante el recorrido de todos los valores posibles de esta variable aleatoria. El gráfico sería similar pero con sesenta saltos muy pequeños. Y si seguimos con un dado de seiscientas caras

$$X = \{0.01; 0.02; 0.03; \dots\dots\dots; 5.97; 5.98; 5.99;\ 6.00\}$$

Si lo miramos desde una pequeña distancia ya no se distinguirán los seiscientos saltos, de 1/600 en probabilidad, de lo que parecería ser un segmento de recta.

Si continuamos el razonamiento con 6000, 60000, 6.000.000 de caras, a los fines prácticos funcionará como si fuera una “pelota”, y si lo llevamos más lejos, podemos imaginar una esfera con “infinitas caras” tomando el límite cuando el número de caras tiende a infinito. Si cada punto de la esfera tiene asociado un número entre 0 y 6, todos tendrán la misma probabilidad de que al rodar se detenga con uno de los puntos hacia arriba (que salga ese número), pero cada número o punto tendrá un infinitésimo de probabilidad.

Nuestro gráfico será más amigable. Si en lugar de los seis saltos del dado cúbico, tenemos los sesenta o seiscientos o seis millones, de saltos cada vez más pequeños, el gráfico se parecerá cada vez más a una línea. Para una esfera será estrictamente un segmento de recta. De modo que el gráfico de la distribución de probabilidad asociada a la esfera tendrá la forma de un segmento de recta como se muestra en la figura que sigue



Nuestra función de distribución de probabilidad acumulada asociada a la esfera numerada en su superficie con los números del cero al seis será un segmento de recta que comienza a acumular probabilidad a partir de *X=0* y habrá acumulado toda la probabilidad posible cuando *X=6*.

Escrito como función de distribución será

$$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x}{6} & 0 \leq x \leq 6 \\ 1 & 6 < x \end{cases}$$

Notemos que la probabilidad acumulada hasta la mitad del rango será

$$P(X \leq 3) = F(3) = \frac{3}{6} = 0{,}5$$

La probabilidad acumulada en la mitad superior del rango será dada por el complemento

$$P(X > 3) = 1 - P(X \leq 3) = 1 - \frac{3}{6} = 0{,}5$$

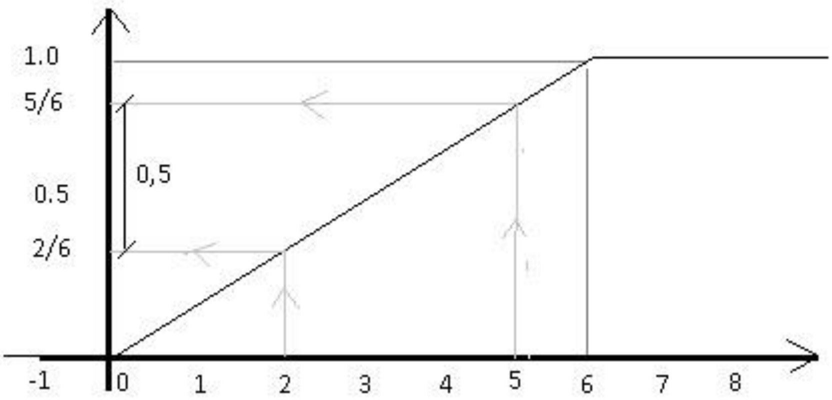
La probabilidad acumulada entre 2 y 5 será la diferencia entre lo acumulado hasta el límite superior indicado (5) y el límite inferior (2).

$$P(2 < X \leq 5) = P(X \leq 5) - P(X \leq 2) = F(5) - F(2) = \frac{5}{6} - \frac{2}{6} = \frac{3}{6} = 0{,}5$$

Tomemos otros límites como $P(1 < X \leq 3) = F(3) - F(1) = \frac{3}{6} - \frac{1}{6} = \frac{2}{6} = 0{,}3333$

En el siguiente gráfico representamos la probabilidad acumulada entre los números 2 y 5. Para ello nos ubicamos en el punto $X = 5$, recorremos el gráfico verticalmente hasta intersectar el segmento de recta y desde allí nos movemos horizontalmente hasta determinar el valor de probabilidad acumulada desde $X = 0$. Obtenemos 5/6. De la misma manera vamos desde $X = 2$ hasta el segmento de recta y desde allí hasta obtener la probabilidad acumulada entre $X = 0$ y

$X = 2$ de valor 2/6. La diferencia entre las dos probabilidades acumuladas nos da como resultado 5/6-2/6=3/6=1/2=0,5, tal como lo habíamos escrito más arriba.



Función de densidad

Definiremos una nueva función como una medida de la rapidez de crecimiento de la función de distribución a medida que crece la variable aleatoria o medida de la velocidad de acumulación de probabilidad. A tal función se la llama "función de densidad" y se nota con letras minúsculas

$$f_{(x)} = función\ de\ densidad$$

El nombre proviene de la noción de densidad aplicada a un cuerpo material. Un cuerpo es más denso cuando más materia contiene en el volumen que ocupa. En el cálculo de probabilidad la "densidad de probabilidad" refiere a cuánta probabilidad se concentra en cada región del dominio en que está definida la variable aleatoria.

Veamos la aplicación concreta a la función de distribución que extrapolamos a nuestro "dado esférico". La velocidad de crecimiento de esta variable aleatoria es constante en el intervalo [0;6] en que la hemos definido, mientras que esa velocidad de crecimiento es nula fuera de ese intervalo. Luego

$$f(x) = \begin{cases} 0 & x < 0 \\ \dfrac{1}{6} & 0 \leq x \leq 6 \\ 0 & 6 < x \end{cases}$$

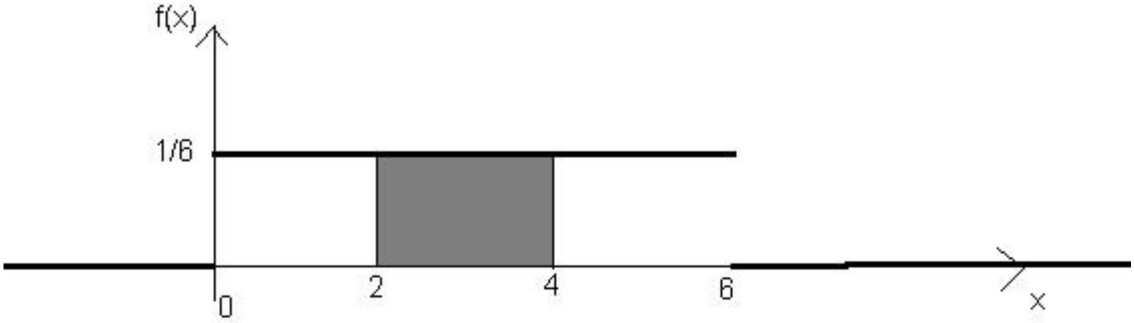
El valor "un sexto" es la pendiente del segmento de recta que crece en la vertical desde cero hasta uno (variable dependiente) cuando en la horizontal la variable aleatoria (variable independiente) se desplaza desde cero hasta seis.

Si representamos gráficamente esta función de densidad se obtiene

Es claro por qué se llama “uniforme” a esta distribución de probabilidad: es constante y vale lo mismo para todos los puntos en los que está definida la variable aleatoria.

Si multiplicamos la base (6) por la altura (1/6) de este rectángulo que queda dibujado por la función de densidad, se obtiene 6*1/6=1. El valor de probabilidad “1” quiere decir que estamos seguros de que algo tiene que ocurrir dentro del intervalo posible y permitido, y es eso lo único de lo que estamos seguros. Por otra parte, en un rectángulo, el producto de la base por la altura nos da el área del rectángulo.

En general, el área encerrada debajo de la función de densidad representa la probabilidad de obtener un resultado dentro del intervalo pedido. Tomemos, a modo de ejemplo $P(2 \leq x \leq 4)$.



El área encerrada debajo de la representación de la función de densidad es un tercio del área total y puede calcularse como “base por altura”, siendo

$$Área = (4-2) * \frac{1}{6} = \frac{2}{6} = \frac{1}{3} = 0{,}333 \ldots$$

Esta distribución de probabilidad es tan usada que hasta tiene un nombre y una nomenclatura, se llama distribución “uniforme”, suele utilizarse para aplicar técnicas de muestreo y caracterizar lo que se llama “ruido aleatorio” en análisis de señales.

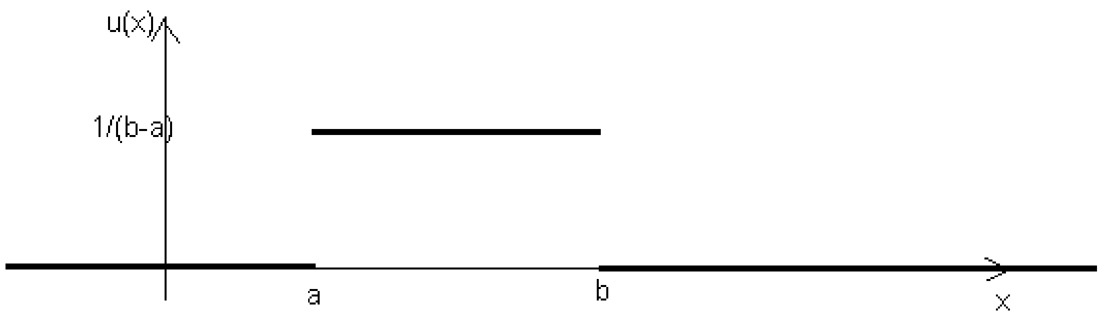
La nomenclatura suele ser $U[a;b]$ para la función de distribución y $u[a;b]$ para la función de densidad definidas

$$u_{[a;b]}(x) = \begin{cases} 0 & x < a \\ \dfrac{1}{b-a} & a \leq x \leq b \\ 0 & b < x \end{cases}$$

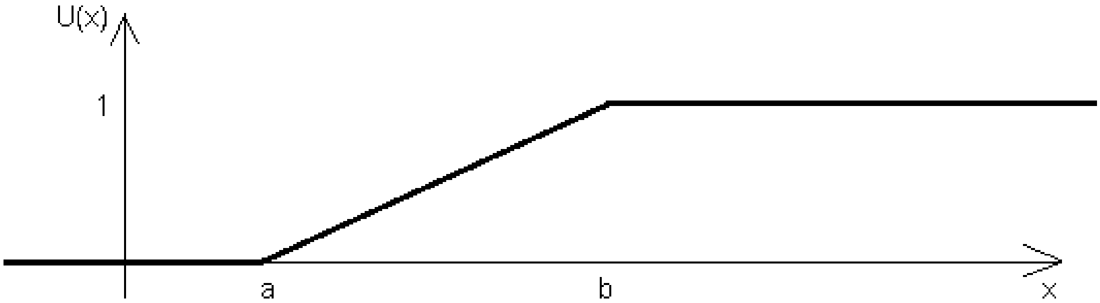
Y

$$U_{[a;b]}(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \leq x \leq b \\ 1 & b < x \end{cases}$$

Con gráficos de densidad



y de distribución



En términos más generales y ya como definición formal, si no hay puntos singulares que tengan una probabilidad finita, la función de distribución será derivable con derivada no negativa. Se definirá la derivada de la función de distribución con el nombre de “función de densidad” de probabilidad

$$f_{(x)} = F'_{(x)}$$

Resulta, por lo tanto

$$F_{(x)} = \int_{-\infty}^{x} f_{(x)} dx$$

Sabemos que la derivada de una función es una medida de su velocidad de crecimiento. Siendo $F_{(x)}$ una función de distribución de probabilidad acumulada, la velocidad de crecimiento dependerá de la “densidad” de probabilidad en cada punto del rango de la variable aleatoria. Recordemos que el término “densidad” se utiliza en relación con la “densidad de masa”, que representa una medida de concentración de masa en cada unidad de volumen. En nuestra aplicación se trata de emular esta noción como una medida de concentración de probabilidad por unidad de variable aleatoria. Suele pensarse la expresión $f_{(x)}dx$ como un elemento diferencial de probabilidad.

Con este planteo resulta

$$P(a < x \leq b) = F_{(b)} - F_{(a)} = \int_{a}^{b} f_{(x)} dx$$

Tomemos la función $f_{(x)} = \frac{3}{8}x^2$ en el intervalo [0,2] y $f_{(x)} = 0$ fuera del intervalo. El gráfico es el de un arco de parábola. Su función de distribución será

$$F_{(x)} = \begin{cases} 0 & x \leq 0 \\ \dfrac{x^3}{8} & 0 < x \leq 2 \\ 1 & 2 < x \end{cases}$$

Puede verse que la probabilidad de obtener un valor en el intervalo central [0,5; 1,5] será $F_{(1,5)} - F_{(0,5)} = \frac{1,5^3}{8} - \frac{0.5^3}{8} = \frac{3,375}{8} - \frac{0,125}{8} = 0,40625$

Es claro también que

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

si se integra entre -∞ y +∞, es decir, sobre todo el eje real, incorporando todos los valores posibles en el cálculo de la probabilidad, es decir que expresa la seguridad absoluta de que “algo” va a ocurrir.

Hemos visto que entre las distribuciones continuas, la más sencilla es la llamada "distribución uniforme" o rectangular. Tal distribución es constante en un intervalo $[a, b]$ con un valor para la densidad

$$f_{(x)} = \frac{1}{b-a} \quad a \le x \le b$$

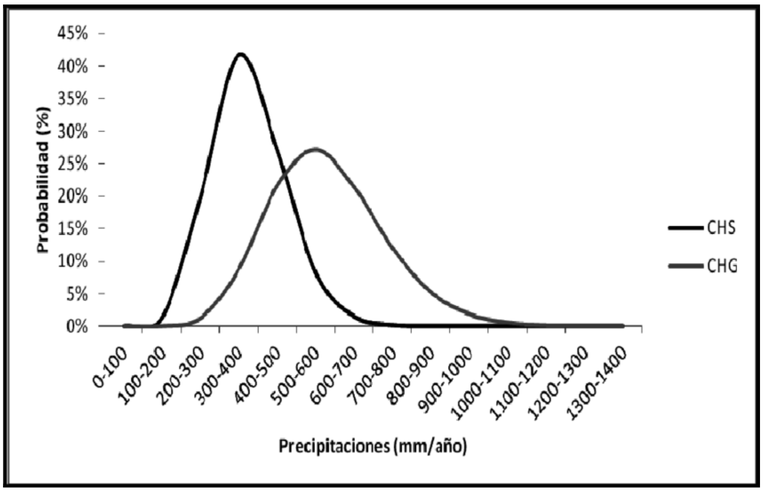
y distribución

$$F_{(x)} = \frac{x-a}{b-a} \quad a \le x \le b$$

Se asume nula la densidad fuera del intervalo, nula la distribución a izquierda de "a" y unitaria a derecha de "b".

Otras funciones de densidad

La que hemos presentado como "distribución uniforme" es la más simple entre las distribuciones continuas y representa la misma probabilidad para todos los puntos del intervalo. Sin embargo las distribuciones pueden tener aspectos muy variados y complejos. Por ejemplo en la siguiente figura presentamos las funciones de densidad de probabilidad de precipitaciones sobre dos ríos de España (Guadalquivir y Segura)

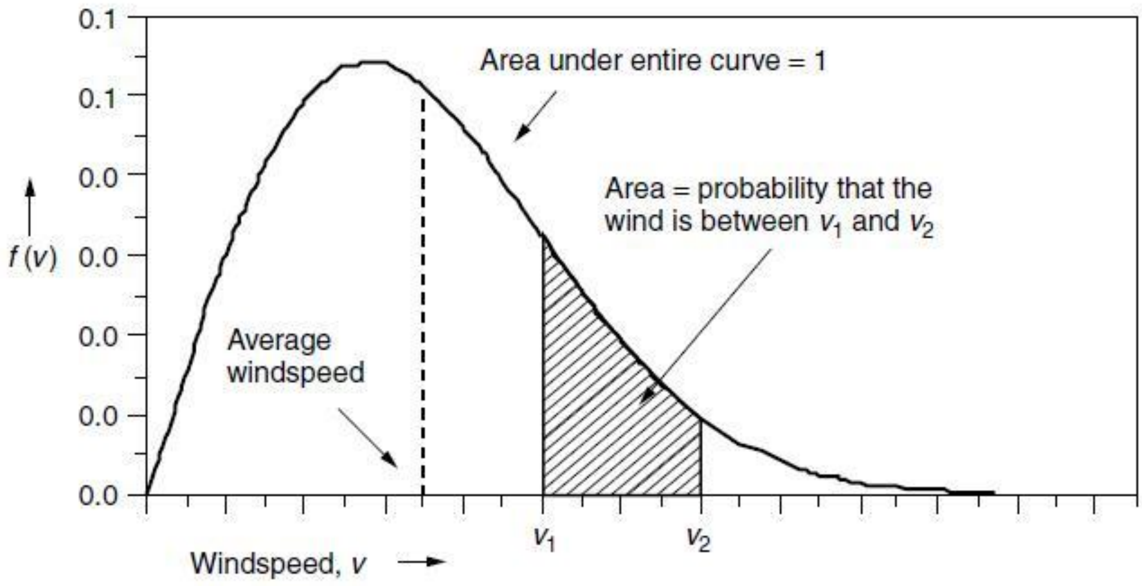


https://www.researchgate.net/figure/Figura-1-Funcion-de-Densidad-de-Probabilidad-para-las-precipitaciones-cuencas-del_fig1_227452493

La curva de la izquierda, del río Segura, tiene un mínimo de precipitación en unos 200mm y un máximo en torno a los 800mm anuales, mientras presenta un máximo de probabilidad en torno a los 400mm (no discutiremos el eje vertical porque corresponde a un análisis estadístico).En cambio la densidad de probabilidad del río Guadalquivir tiene un mínimo de precipitación en torno a los 300mm y un máximo en el orden de 1200mm, con un máximo de probabilidad alrededor de 650mm. La distribución del río Segura es más concentrada que la del río Guadalquivir.

El área encerrada debajo de ambas curvas vale uno, de modo que si se quiere calcular la probabilidad de que la precipitación sobre el río Segura sea menor que 500mm bastaría calcular el área debajo de la curva que le corresponde a la izquierda de un segmento que corte el eje horizontal en 500mm. Una observación del gráfico permite estimar esta probabilidad en el orden del 60%, mientras que la misma probabilidad estimada para el río Guadalquivir difícilmente alcance el 20%.

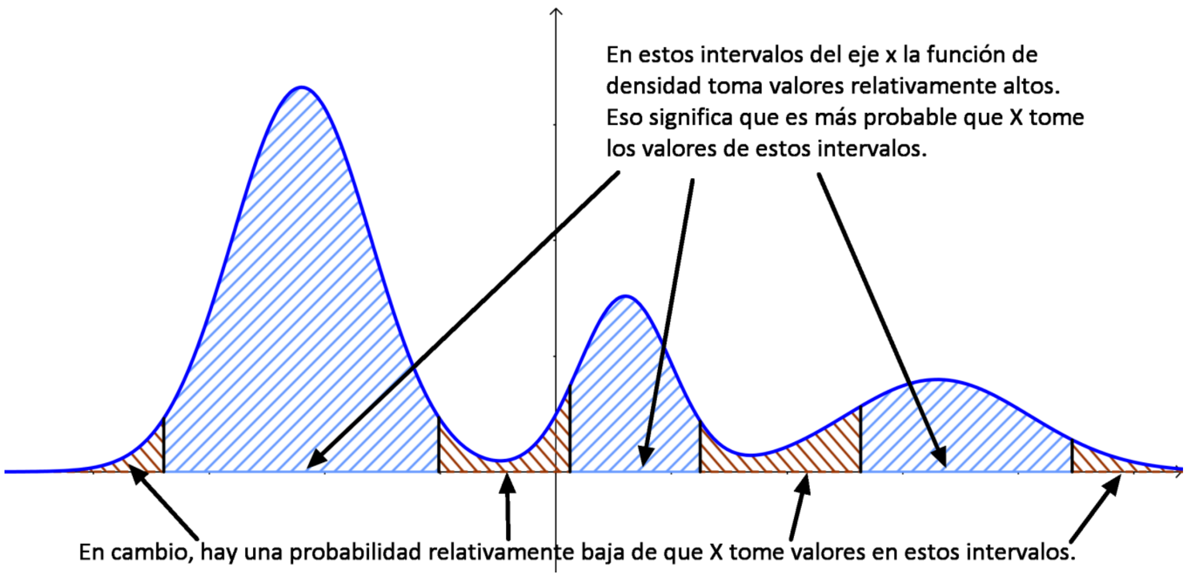
El siguiente gráfico es una simulación de una forma típica de la densidad de probabilidad de la velocidad del viento para estudios de energía eólica



A windspeed probability density function (p.d.f).

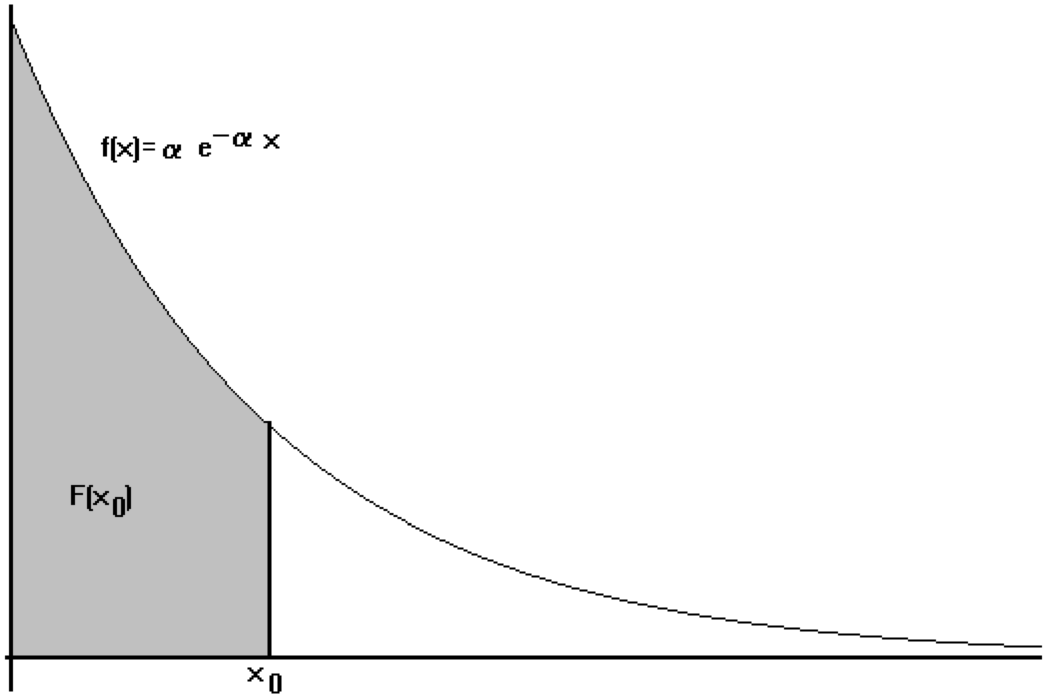
Luego veremos el significado de la línea de trazos (Average windspeed o velocidad promedio del viento). Por el momento observemos el área sombreada que representa la probabilidad de que la velocidad del viento se encuentre entre V_1 y V_2.

La siguiente es una función de densidad que representa intervalos con altos valores de probabilidad y otras regiones con mínimos de probabilidad. Puede ser comparable a distribuciones de probabilidad de detección de un electrón en las proximidades del núcleo de un átomo como el hidrógeno o el helio.



https://postdata-statistics.com/introestadistica/impartidos/quimicauah20142015/sesion05b/sesion05b

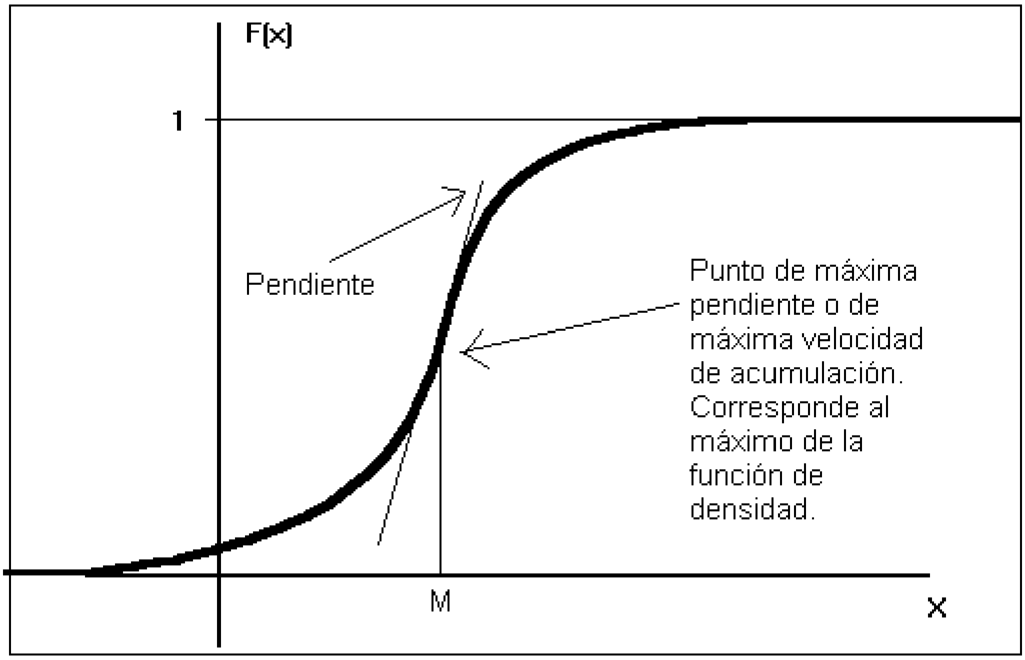
La siguiente tiene un nombre, se llama “función de densidad exponencial”, y tiene aplicaciones en tiempos de espera y procesos de decaimiento radioactivo. Notemos el interés puesto en el cálculo del área como probabilidad de que el resultado sea menor que X_0



https://www.uv.es/ceaces/base/modelos%20de%20probabilidad/MODEPR1.htm

Vemos que hay variadas formas de las funciones de densidad que describen el comportamiento probabilista de muchos sistemas reales. El cálculo de probabilidades se realiza por medio del cálculo de áreas encerradas bajo las curvas, o bien si se dispone de las funciones de distribución.

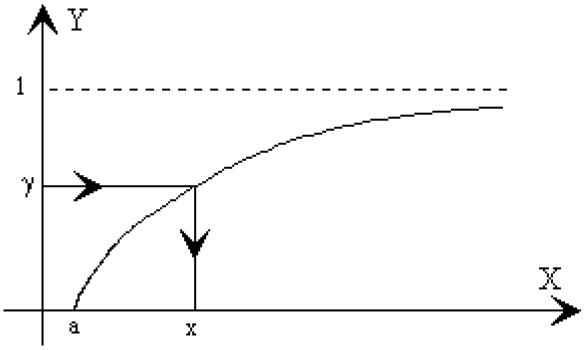
Una forma típica de las funciones de distribución correspondientes a los primeros gráficos es



https://www.uv.es/ceaces/base/variable%20aleatoria/varalea.htm

El punto de máxima inclinación de la función de distribución o punto de máxima velocidad de acumulación corresponde al punto de máxima densidad de probabilidad. Si presentan tres máximos de probabilidad la distribución será sinuosa con tres máximos de inclinación.

La siguiente función de distribución es típica de la correspondiente a la función de densidad exponencial. En el ejemplo gráfico, a un valor de la variable aleatoria “x” le corresponde una probabilidad acumulada “y”.



http://www.sc.ehu.es/sbweb/fisica/cursoJava/numerico/montecarlo/aleatoria/aleatoria.htm

En general se utiliza funciones explícitas cuando es posible, tablas de distribución de probabilidad acumulada cuando éstas están disponibles, o bien una estimación gráfica cuando no hay métodos más precisos de cálculo.

Cambio de variables

En ocasiones puede interesar la distribución de probabilidad de una variable aleatoria pero se dispone de la distribución de otra variable aleatoria, relacionada con la de interés. Por ejemplo, se conoce la distribución de probabilidad de producción de algún alimento o de extracción de un mineral, pero se tiene interés en la distribución de probabilidad de tiempos, costos o beneficios asociados con la producción. Si existe alguna función que relacione en forma determinista una variable con otro, puede hacerse un cambio de variables en la definición de las distribuciones.

Supongamos que $Y_{(X)}$ es una función biyectiva monótona creciente que relaciona la variable aleatoria X con otra Y, y que por lo tanto existe $X^{-1}_{(Y)}$, su inversa, también monótona creciente. Notemos $G_{(y)}$ como la función de distribución de la variable Y. Como por definición

$$G_{(y)} = P(Y \leq y) = P\left(Y_{(X)} \leq y\right) = P\left(X \leq X^{-1}_{(y)}\right) = F_{\left(X^{-1}_{(y)}\right)}$$

cuya derivada devuelve la función de densidad de Y,

$$g_{(y)} = \frac{dF_{\left(X^{-1}_{(y)}\right)}}{dy}$$

Por ejemplo, si $F_{(x)} = \frac{x-3}{5-3}$, uniforme en el intervalo [3,5], e $Y = 4X - 1$, resulta que $G_{(y)} = P(Y \leq y) = P(4X - 1 \leq y) = P\left(X \leq \frac{y+1}{4}\right) = F_{\left(\frac{(y+1)}{4}\right)} = \frac{\frac{y+1}{4}-3}{5-3} = \frac{y-11}{8}$, cuya derivada vale $g_{(y)} = \frac{1}{8}$, definida en el intervalo [11, 19], es decir que una uniforme en el intervalo [3,5] de valor constante 0,5, se traduce en una uniforme en el intervalo [11, 19] con un valor constante de 0,125.

Si la función fuese monótona decreciente, basta invertir el signo en la desigualdad o usar el complemento, y si no es biyectiva, basta separar el dominio en subdominios de biyectividad o restringir el codominio (por ejemplo, tomar la parte positiva de una raíz cuadrada).

Probabilidades y Estadística: Ejercicios Unidad Nº 3

Variables Aleatorias y Distribuciones

1. Dados los siguientes experimentos, definir una variable aleatoria asociada a cada uno y caracterizar el tipo de variable aleatoria.
a) Tirar una moneda.
b) Tirar un dado.
c) Hacer cien disparos a un blanco.
d) El salto en largo de un atleta.
e) La temperatura de Campana.
f) Formación de dúos musicales mixtos con 3 mujeres y 2 hombres postulantes.

2. En el punto f del ejercicio anterior la selección de los integrantes del dúo se hace al azar, por lo cual el dúo puede formarse por dos hombres, dos mujeres, o resultar mixto. Hallar la distribución de probabilidad asociada al evento contabilizando el número de mujeres que integran el dúo.

3. Sea X una variable aleatoria distribuida binomialmente con base en 10 repeticiones de un experimento. Si $p = 0{,}3$, calcular las siguientes probabilidades:

$$a)P(X \leq 8) \qquad b)P(X = 7) \qquad c)P(X > 6)$$

4. Tres de cada diez automóviles doblan a la derecha en una determinada esquina. Calcular la probabilidad de que al menos uno de cinco automóviles determinados doble a la derecha y la probabilidad de que exactamente uno de los cinco vehículos doble a la derecha.

5. Se sabe que al tirar una moneda sale cara tres veces más a menudo que ceca. Esta moneda se lanza tres veces. Sea X el número de caras que aparecen. Establecer la distribución de probabilidades de X y también la distribución acumulada. Hacer un gráfico de la distribución puntual y acumulada.

6. De un lote que contiene 25 artículos, 5 de los cuales son defectuosos, se eligen 4 al azar. Sea X el número de artículos defectuosos encontrados. Obtener la distribución de probabilidad de X si a) los artículos se escogen con reposición, b) los artículos se escogen sin reposición. Hacer en cada caso un gráfico de la distribución puntual y acumulada.

7. La variable aleatoria continua X tiene la función de densidad $f_{(x)} = 3x^2$ para $0 \leq x \leq 1$. Dibujarla la función de densidad, hallar la función de distribución y dibujarla. a) Calcular la probabilidad de que un resultado se encuentre en el intervalo (0,2; 0,5). b) Hallar la probabilidad de que el resultado sea mayor que 0,5. c) Hallar la probabilidad de que el resultado sea menor que 0,2.

8. Una distribución definida sobre el intervalo [0,1] tiene la forma ax^3 para cierto valor de a. Dibujarla cualitativamente, hallar la función de distribución y graficarla cualitativamente. a) Hallar el valor de a. b) Hallar la probabilidad de que el resultado del experimento sea un valor menor que 0.7 sabiendo que se obtuvo un valor mayor que 0.5 c) Si z es un número que satisface $0 < z < 1$, calcular $P(X > Z/X > z/2)$.

9. Se supone que el diámetro de un cable eléctrico, representado por X, es una variable aleatoria continua con función de densidad $f_{(x)} = -6x(x-1)$ para $0 \leq x \leq 1$.
a) Verificar que la anterior es una función de densidad y dibujarla.
b) Obtener una expresión para la función de distribución de X y dibujarla.
c) Calcular $P(X \leq ½)$ utilizando la función de densidad.

d) Calcular $P(X \leq ½)$ utilizando la función de distribución.
e) Calcular $P(X \leq ½/1/3 < x < 2/3)$ utilizando las funciones de densidad y distribución.

10. La variable aleatoria continua X tiene función de densidad de probabilidad dad por $f_{(x)} = \frac{x}{2}$ para $0 \leq x \leq 2$. Se hacen dos determinaciones independientes de X. ¿Cuánto vale la probabilidad de que ambas determinaciones sean mayores que 1? Si se han hecho tres determinaciones independientes, ¿cuánto vale la probabilidad de que exactamente dos sean mayores que 1? ¿Cuánto vale la probabilidad de que al menos dos determinaciones sean mayores que 1? Si se repite el experimento 8 veces, hallar la probabilidad de que dos determinaciones sean mayores que 1. Si se repite el experimento quince veces, hallar la probabilidad de que al menos una determinación sea mayor que 1.

11. Calcular la probabilidad de que con una distribución uniforme en el intervalo [0,10] se obtenga un resultado en el intervalo [3,8]. Hallar esta probabilidad sabiendo que el resultado obtenido es estrictamente mayor que 5. ¿Cuánto vale $P(X = 5)$ en la distribución uniforme continua?

12. Suponer que la temperatura de reacción X (en°C) de un proceso químico tiene una distribución uniforme con extremos A = -5 y B = 5.
a) Calcular $P(X \leq 0)$ b) Calcular $P(-2 < X \leq 2)$
c) Calcular $P(-2 < X \leq 3)$ d) Calcular $P(X \leq 0/X \leq 3)$

13. Supóngase que X está distribuida uniformemente en [-a,a] con a>0. Cada vez que sea posible, determinar "a" de modo que satisfaga lo siguiente:
a) $P(X > 1) = \frac{1}{3}$ b) $P(X > 1) = \frac{1}{2}$ c) $P\left(X \leq \frac{1}{2}\right) = 0{,}7$ d) $P\left(X \leq \frac{1}{2}\right) = 0{,}3$
e) $P(|X| \leq 1) = P(|X| > 1)$

14. La función de densidad dada por $f(x) = x/4$ está definida en el intervalo (1, 3). Hallar la función de distribución de $f(x)$. Si $Y = 2X + 3$, hallar las funciones de distribución y de densidad de Y.

15. Una corriente eléctrica I, medida en amperes, que fluctúa al azar, se puede considerar como una variable aleatoria distribuida uniformemente en el intervalo real (9; 11). Si esta corriente pasa por un conductor con resistencia de R=2 medida en ohm, encontrar la función de densidad de probabilidad de la potencia, sabiendo que se calcula de acuerdo con la ecuación $P = RI^2$.

Ejercicios resueltos

1. Dados los siguientes experimentos, definir una variable aleatoria asociada a cada uno y caracterizar el tipo de variable aleatoria.

a) Tirar una moneda.
b) Tirar un dado.
c) Hacer cien disparos a un blanco.
d) El salto en largo de un atleta.
e) La temperatura de Campana.
f) Formación de dúos musicales mixtos con 3 mujeres y 2 hombres postulantes.

a) Si se asume que el experimento tiene los resultados "cara" y "ceca", es discreta.
b) Si se asume que resulta una de las caras, es discreta.
c) Si contabilizamos el número de blancos, es discreta; si medimos la distancia al blanco, es continua.
d) El salto en largo se mide en distancia y es continua.
e) La temperatura es una variable aleatoria continua por definición.
f) Podemos especificar el número de mujeres que lo componen y resulta discreta.

2. En el punto f del ejercicio anterior la selección de los integrantes del dúo se hace al azar, por lo cual el dúo puede formarse por dos hombres, dos mujeres, o resultar mixto. Hallar la distribución de probabilidad asociada al evento contabilizando el número de mujeres que integran el dúo.

Si la variable aleatoria es el número de mujeres que integran el dúo, éste puede estar formado por dos hombres (X=0), ser mixto (X=1) o por dos mujeres (X=2). Podemos plantearlo como una distribución hipergeométrica o ser más prácticos y contabilizar diez dúos posibles como una combinación de dos elementos tomados entre cinco distinguibles. En sólo un caso (los dos hombres) X=0, hay 3*2=6 opciones para que el dúo sea mixto y tres formas de armar un dúo con las tres mujeres. Luego

$$P(X=0)=0{,}1 \qquad P(X=1)=0{,}6 \qquad P(X=2)=0{,}3$$
$$P(X\leq 0)=0{,}1 \qquad P(X\leq 1)=0{,}7 \qquad P(X\leq 2)=1{,}0$$

Con estos datos construimos los gráficos de distribución puntual y acumulada.

3. Sea X una variable aleatoria distribuida binomialmente con base en 10 repeticiones de un experimento. Si $p=0{,}3$, calcular las siguientes probabilidades:

$$a)P(X\leq 8) \qquad b)P(X=7) \qquad c)P(X>6)$$

El objetivo del ejercicio es sólo practicar el uso de la distribución de Bernoulli. Para no hallar todas las probabilidades desde X=0 hasta X=8 podemos calcular

$$P(X=9)=\binom{10}{9}0{,}3^{9}(1-0{,}3)^{10-9}=0.000137781$$
$$P(X=10)=\binom{10}{10}0{,}3^{10}(1-0{,}3)^{10-10}=0.0000059049$$

Luego $P(X\leq 8)=1-P(X>8)=1-P(X=9)-P(X=10)=0.9998563141$

$$P(X=7)=\binom{10}{7}0{,}3^{7}(1-0{,}3)^{10-7}=0.009001692$$

Y como

$$P(X=8)=\binom{10}{8}0{,}3^{8}(1-0{,}3)^{10-8}=0.0014460005$$

Y $P(X>6)=0{,}0105920784$

4. Tres de cada diez automóviles doblan a la derecha en una determinada esquina. Calcular la probabilidad de que al menos uno de cinco automóviles determinados doble a la derecha y la probabilidad de que exactamente uno de los cinco vehículos doble a la derecha.

Asumimos que los automóviles no se persiguen entre sí ni van en caravana, por lo tanto que se mueven independientemente. Si tres de cada diez doblan a la derecha, asumimos que 0,3 es la probabilidad de que un automóvil cualquiera doble a la derecha. La probabilidad de que al menos uno doble a la derecha se obtiene a partir del complemento del evento "que no doble ninguno a la derecha". Luego

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{5}{0} 0{,}3^0 (1 - 0{,}3)^{5-0} = 0{,}83193$$

La probabilidad de que exactamente uno doble a la derecha se obtiene de

$$P(X = 1) = \binom{5}{1} 0{,}3^1 (1 - 0{,}3)^{5-1} = 0{,}36015$$

5. Se sabe que al tirar una moneda sale cara tres veces más a menudo que ceca. Esta moneda se lanza tres veces. Sea X el número de caras que aparecen. Establecer la función de distribución de probabilidades de X y también la función de distribución acumulada. Hacer un gráfico de la distribución puntual y acumulada.

Dado que P(cara)=3P(ceca) y P(cara)+P(ceca)=1, resulta P(cara)=0,75 y P(ceca)=0,25.
Como se lanza tres veces X={0;1;2;3}

$$P(X = 0) = \binom{3}{0} 0{,}75^0 (1 - 0{,}75)^{3-0} = 0{,}015625$$
$$P(X = 1) = \binom{3}{1} 0{,}75^1 (1 - 0{,}75)^{3-1} = 0{,}140625$$
$$P(X = 2) = \binom{3}{2} 0{,}75^2 (1 - 0{,}75)^{3-2} = 0{,}421875$$
$$P(X = 3) = \binom{3}{3} 0{,}75^3 (1 - 0{,}75)^{3-3} = 0{,}421875$$

Tenemos entonces la distribución puntual, y para la acumulada

$$P(X \leq 0) = 0{,}015625$$
$$P(X \leq 1) = 0{,}15625$$
$$P(X \leq 2) = 0{,}578125$$
$$P(X \leq 3) = 1{,}00$$

6. De un lote que contiene 25 artículos, 5 de los cuales son defectuosos, se eligen 4 al azar. Sea X el número de artículos defectuosos encontrados. Obtener la distribución de probabilidad de X si a) los artículos se escogen con reposición, b) los artículos se escogen sin reposición. Hacer en cada caso un gráfico de la distribución puntual y acumulada.

En el punto (a), se es escogen con reposición, tenemos una distribución binomial dado que el espacio muestral es siempre el mismo y las extracciones son independientes. La probabilidad de que un artículo sea defectuoso vale P(Defectuoso)=5/25=0,2. Resulta

$$P(X = k) = \binom{4}{k} 0{,}2^k (1 - 0{,}2)^{4-k}$$

Si se escogen sin reposición tenemos la aplicación de la distribución hipergeométrica

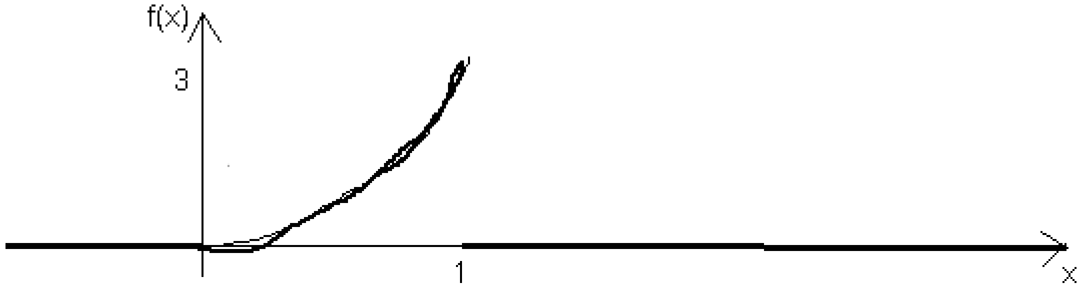
$$P(X = k) = \frac{\binom{5}{k}\binom{20}{4-k}}{\binom{5+20}{4}}$$

La resolución para los distintos valores de "k" entre k=0 y k=4 dará la distribución puntual en cada caso y, la suma progresiva, la distribución acumulada.

7. La variable aleatoria continua X tiene la función de densidad $f_{(x)} = 3x^2$ para $0 \leq x \leq 1$. Dibujarla la función de densidad, hallar la función de distribución y dibujarla. a) Calcular la probabilidad de que un resultado se encuentre en el intervalo (0,2; 0,5). b) Hallar la probabilidad de que el resultado sea mayor que 0,5. c) Hallar la probabilidad de que el resultado sea menor que 0,2.

Partiendo de la definición

$$f(x) = \begin{cases} 0 & x < 0 \\ 3x^2 & 0 \leq x \leq 1 \\ 0 & 1 < x \end{cases}$$

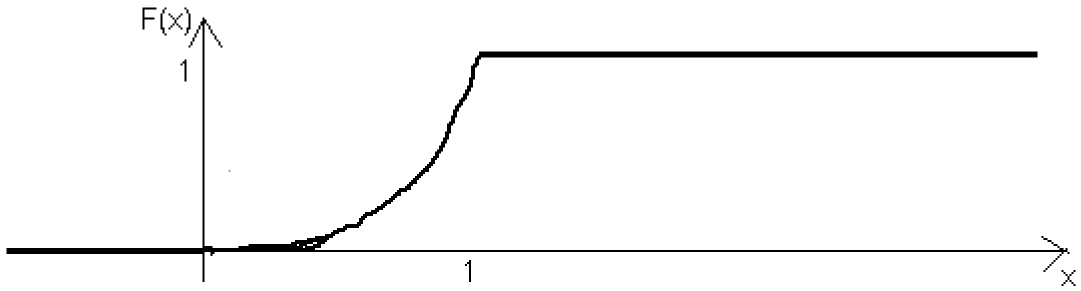


Para hallar la función de distribución basta integrar

$$F(x) = \int_{-\infty}^{x} f(x)dx = \int_{0}^{x} 3x^2 dx = x^3$$

Con lo cual

$$F(x) = \begin{cases} 0 & < 0 \\ x^3 & 0 \leq x \leq 1 \\ 1 & 1 < x \end{cases}$$



Más allá de la calidad de los gráficos, se pretende destacar las características de las representaciones gráficas de las funciones de densidad y de distribución.
Para el cálculo de probabilidad usando la función de densidad recurrimos a la integración y a la regla de Barrow a partir de la primitiva

$$P(0{,}2 < x \leq 0{,}5) = \int_{0{,}2}^{0{,}5} 3x^2 dx = 0{,}5^3 - 0{,}2^3 = 0{,}125 - 0.08 = 0{,}117 = 11{,}7\%$$

Gráficamente es representada por el área bajo la curva en el gráfico de f(x) en el dominio de integración.
Para el cálculo de probabilidad usando la función de distribución recurrimos a la definición y a la diferencia entre probabilidades acumuladas

$$P(0{,}2 < x \leq 0{,}5) = P(x \leq 0{,}5) - P(x \leq 0{,}2) = F(0{,}5) - F(0{,}2) = 0{,}117$$

En el gráfico de la función de distribución, es representada por la diferencia entre las imágenes de 0,5 y de 0,2 en el eje vertical.
Abreviando el comentario, para la probabilidad de obtener un resultado mayor que 0,5 usando la función de densidad

$$P(x > 0{,}5) = \int_{0{,}5}^{\infty} f(x)dx = \int_{0{,}5}^{1} 3x^2 dx = x^3|_{0{,}5}^{1} = 1^3 - 0{,}5^3 = 1 - 0{,}125 = 0{,}875$$

Y usando la de distribución

$$P(x > 0{,}5) = F(+\infty) - F(0{,}5) = 1 - P(x \leq 0{,}5) = 1 - F(0{,}5) = 1 - 0{,}5^3 = 0{,}875$$

Hemos usado variantes en la notación del cálculo pero respetado la interpretación del planteo en el recurso a la función de densidad a través de una integración en un intervalo, y de la diferencia entre dos valores de probabilidad acumulada al usar la distribución.
Finalmente

$$P(x < 0{,}2) = \int_0^{0{,}2} 3x^2 dx = P(x \leq 0{,}2) = F(0{,}2) = 0{,}2^3 = 0{,}08 = 8\%$$

Incorporamos los dos recursos en un solo desarrollo. Hacemos notar que formalmente debemos incluir el 0,2 en la desigualdad para usar la función de distribución, pero en el cálculo es irrelevante incorporar un punto en una integración, a menos que haya una discontinuidad en ese punto.

8. Una distribución definida sobre el intervalo [0,1] tiene la forma ax^3 para cierto valor de a. Dibujarla cualitativamente, hallar la función de distribución y graficarla cualitativamente. a) Hallar el valor de a. b) Hallar la probabilidad de que el resultado del experimento sea un valor menor que 0.7 sabiendo que se obtuvo un valor mayor que 0.5 c) Si *z* es un número que satisface $0 < z < 1$, calcular $P(X > Z/X > z/2)$.

El planteo del ejercicio 8 incorpora un ajuste de escala a la determinación de las funciones de densidad y de distribución a través del parámetro "*a*". Sabemos que debe ser

$$\int_{-\infty}^{+\infty} f(x)dx = \int_0^1 ax^3 dx = 1$$

Luego

$$a \left. \frac{x^4}{4} \right|_0^1 = a\left(\frac{1^4}{4} - \frac{0^4}{4}\right) = \frac{a}{4} = 1 \quad \to a = 4$$

Luego

$$f(x) = 4x^3 \ \ si \quad 0 \leq x \leq 1 \qquad y \qquad f(x) = 0 \quad si \quad x \sim \in [0;1]$$

Y

$$F(x) = x^4 \quad si \quad x \in [0;1] \qquad y \qquad F(x) = 0 \quad si \quad x < 0 \qquad y \qquad F(x) = 1 \quad si \quad x > 1$$

El siguiente punto es una probabilidad condicional.

$$P(x < 0{,}7/x > 0{,}5) = \frac{P(x \in [0;0{,}7) \cap (0{,}5;1])}{P(x \in (0{,}5;1])} = \frac{P\big(x \in (0{,}5;0{,}7)\big)}{P(x \in (0{,}5;1])} = \frac{\int_{0{,}5}^{0{,}7} 4x^3 dx}{\int_{0{,}5}^{1} 4x^3 dx} \cong 19\%$$

El punto final también es una probabilidad condicional paramétrica

$$P\left(x > z/x > \frac{z}{2}\right) = \frac{P\left(x \in (z;1] \cap x \in (\frac{z}{2};1])\right)}{P\left(x \in (\frac{z}{2};1])\right)} = \frac{P(x \in (z;1])}{P\left(x \in (\frac{z}{2};1])\right)} = \frac{1 - z^4}{1 - \left(\frac{z}{2}\right)^4}$$

Si planteamos, como punto adicional

$$P\left(x \leq \frac{z}{2}/x \leq z\right) = \frac{P\left(x \in \left[0;\frac{z}{2}\right] \cap x \in [0;z])\right)}{P(x \in [0;z]))} = \frac{P\left(x \in \left[0;\frac{z}{2}\right]\right)}{P(x \in [0;z])} = \frac{\left(\frac{z}{2}\right)^4}{z^4} = \frac{1}{16}$$

9. Se supone que el diámetro de un cable eléctrico, representado por *X*, es una variable aleatoria continua con función de densidad $f_{(x)} = -6x(x-1)$ para $0 \leq x \leq 1$.
   a) Verificar que la anterior es una función de densidad y dibujarla.
   b) Obtener una expresión para la función de distribución de *X* y dibujarla.
   c) Calcular $P(X \leq ½)$ utilizando la función de densidad.
   d) Calcular $P(X \leq ½)$ utilizando la función de distribución.
   e) Calcular $P(X \leq ½/1/3 < x < 2/3)$ utilizando las funciones de densidad y distribución.

Por el signo que precede a la distribución parece ser negativa, pero es un arco de parábola que tiene concavidad negativa, raíces en $x = 0$ y en $x = 1$ con un máximo positivo en el punto $x = 0{,}5$. Es positiva en todo el intervalo y basta verificar

$$-6\int_0^1 x(x-1)dx = -6\left(\int_0^1 x^2dx - \int_0^1 xdx\right) = -6\left(\frac{1}{3}-\frac{1}{2}\right) = -6\left(-\frac{1}{6}\right) = 1$$

Para obtener la función de distribución

$$F(x) = -6\left(\int_0^x x^2dx - \int_0^x xdx\right) = -6\left(\frac{x^3}{3}-\frac{x^2}{2}\right) = 3x^2 - 2x^3 \quad si\ 0 \le x \le 1$$

$$P\left(x \le \frac{1}{2}\right) = \int_0^{0,5} x(x-1)dx = 0{,}5$$

$$P\left(x \le \frac{1}{2}\right) = F(0{,}5) = 3 * 0{,}5^2 - 2 * 0{,}5^3$$

$$P\left(x \in \left[0;\frac{1}{2}\right] / x \in \left(\frac{1}{3};\frac{2}{3}\right)\right) = \frac{P\left(x \in \left[0;\frac{1}{2}\right] \cap x \in \left(\frac{1}{3};\frac{2}{3}\right)\right)}{P\left(x \in \left(\frac{1}{3};\frac{2}{3}\right)\right)} = \frac{P\left(x \in \left(\frac{1}{3};\frac{1}{2}\right]\right)}{P\left(x \in \left(\frac{1}{3};\frac{2}{3}\right)\right)} =$$

$$= \frac{\left(3\left(\frac{1}{2}\right)^2 - 2\left(\frac{1}{2}\right)^3\right) - \left(3\left(\frac{1}{3}\right)^2 - 2\left(\frac{1}{3}\right)^3\right)}{\left(3\left(\frac{2}{3}\right)^2 - 2\left(\frac{2}{3}\right)^3\right) - \left(3\left(\frac{1}{3}\right)^2 - 2\left(\frac{1}{3}\right)^3\right)} = \frac{\frac{3}{4}-\frac{2}{8}-\left(\frac{3}{9}-\frac{2}{27}\right)}{\frac{12}{9}-\frac{16}{27}-\left(\frac{3}{9}-\frac{2}{27}\right)} =$$

$$= \frac{\frac{1}{2}-\frac{7}{27}}{1-\frac{14}{27}} = \frac{\frac{27-14}{54}}{\frac{27-14}{27}} = \frac{1}{2} = 0{,}5 = 50\%$$

10. La variable aleatoria continua *X* tiene función de densidad de probabilidad dad por $f_{(x)} = \frac{x}{2}$ para $0 \le x \le 2$. Se hacen dos determinaciones independientes de *X*. ¿Cuánto vale la probabilidad de que ambas determinaciones sean mayores que 1? Si se han hecho tres determinaciones independientes, ¿cuánto vale la probabilidad de que exactamente dos sean mayores que 1? ¿Cuánto vale la probabilidad de que al menos dos determinaciones sean mayores que 1? Si se repite el experimento 8 veces, hallar la probabilidad de que dos determinaciones sean mayores que 1. Si se repite el experimento quince veces, hallar la probabilidad de que al menos una determinación sea mayor que 1.

Para calcular la probabilidad de que ambas determinaciones sean mayores que uno hay que comenzar calculando la probabilidad de que una de ellas sea mayor que uno, luego

$$P(x > 1) = \int_1^2 \frac{x}{2}dx = \left.\frac{x^2}{4}\right|_1^2 = 1 - \frac{1}{4} = 0{,}75$$

Como son dos determinaciones independientes

$$P(x_1 > 1 \wedge x_2 > 1) = P(x > 1)^2 = 0{,}75^2 = \frac{9}{16} = 0{,}5625$$

Para determinar la probabilidad de que dos de tres determinaciones independientes de la misma variable aleatoria sean mayores que uno, vemos que estamos en las condiciones de recurrir a la distribución binomial. Si llamamos “k” al número de determinaciones mayores que uno y ponemos un subíndice para indicar el número de repeticiones

$$P_3(k=2)=\binom{3}{2}0{,}75^2(1-0{,}75)^{3-2}=3*0{,}5625*0.25=0{,}421875\cong 42{,}2\%$$

Si se repite ocho veces, hallar la probabilidad de que dos determinaciones sean mayores que uno nos lleva al mismo planteo con parámetros diferentes en la binomial

$$P_8(k=2)=\binom{8}{2}0{,}75^2(1-0{,}75)^{8-2}=0{,}003845\cong 0{,}38\%$$

Si se repite quince veces, para hallar la probabilidad de que al menos una determinación sea mayor que uno, recurrimos al complemento e implícitamente a la binomial.

$$P_{15}(k>1)=1-P_{15}(k=0)=1-(1-0{,}75)^{15}\cong 1$$

11. Calcular la probabilidad de que con una distribución uniforme en el intervalo [0,10] se obtenga un resultado en el intervalo [3,8]. Hallar esta probabilidad sabiendo que el resultado obtenido es estrictamente mayor que 5. ¿Cuánto vale $P(X=5)$ en la distribución uniforme continua?

Si recordamos que una distribución uniforme es constante en el intervalo en que está definida y equivalente a la recíproca del tamaño de su tamaño, dado que el intervalo es el rango [0;10], la constante debe ser 1/10=0,1.

Para hallar la probabilidad asociada al intervalo [3;8] formalmente integramos

$$P(3\le x\le 8)=\int_3^8 0{,}1dx=0{,}1*x|_3^8=0{,}1*(8-3)=0{,}5$$

La integración es inmediata y nos hemos tomado el trabajo de recurrir al teorema de Barrow, recurramos a la definición formal de la función de distribución

$$P(3\le x\le 8)=P(X=3)+P(3<x\le 8)=0+U(8)-U(3)=\frac{8-0}{10-0}-\frac{3-0}{10-0}=0{,}5$$

Si bien el resultado es el mismo y el procedimiento también es correcto, el uso de Barrow es un recurso de cálculo y el segundo es un procedimiento formal para el cual se tuvo que hacer una separación para $X=3$ y $X\in(3;8]$. Si la distribución no fuese continua en $x=3$, los procedimientos no habrían sido equivalentes y no se hubiese podido integrar utilizando el teorema de Barrow.

La segunda parte propone una probabilidad condicional

$$P(3\le x\le 8/5\le x\le 10)=\frac{P(x\in[3;8]\cap[5;10])}{P(x\in[5;10])}=\frac{P(5\le x\le 8)}{P(5\le x\le 10)}=\frac{0{,}3}{0{,}5}=0{,}6$$

La última pregunta nos remite al problema de la igualdad inclusiva o la igualdad estricta al plantear específicamente P(X=5). Desde el punto de vista del cálculo, P(X=5) nos obligaría a integrar una función continua entre $x=5$ y $x=5$, lo cual da cero. Y aplicando la función de distribución

$$P(X=5)=U(5)-\lim_{x\to 5^-}U(x)=0{,}5-\lim_{x\to 5^-}\frac{x}{10}=0{,}5-0{,}5=0$$

Recordemos que esto no quiere decir que es imposible que el resultado sea exactamente $X=5$ sino que no es posible pretender un conocimiento predictivo de la factibilidad de este resultado previamente a la realización del experimento. Por otra parte, si pasamos por alto la incorporación del límite inferior en el primer punto del ejercicio, éste resultado lo hace evidente.

12. Suponer que la temperatura de reacción X (en °C) de un proceso químico tiene una distribución uniforme con extremos A = -5 y B = 5.
   a) Calcular $P(X \leq 0)$ b) Calcular $P(-2 < X \leq 2)$
   c) Calcular $P(-2 < X \leq 3)$ d) Calcular $P(X \leq 0/X \leq 3)$

Omitiremos las unidades. Dado que es uniforme en el intervalo [-5;5], el tamaño del intervalo vale TI=5-(-5)=10. Sobre esta base la constante de integración de la distribución uniforme vale 0,1, luego

$$P(X \leq 0) = F(0) = \frac{0-(-5)}{5-(-5)} = \int_{-5}^{0} 0{,}1dx = 0{,}5$$
$$P(-2 < X \leq 2) = F(2) - F(-2) = \frac{2-(-5)}{5-(-5)} - \frac{-2-(-5)}{5-(-5)} = \int_{-2}^{2} 0{,}1dx = 0{,}4$$
$$P(-2 < X \leq 3) = F(3) - F(-2) = \frac{3-(-5)}{5-(-5)} - \frac{-2-(-5)}{5-(-5)} = \int_{-2}^{3} 0{,}1dx = 0{,}5$$
$$P(X \leq 0/X \leq 3) = \frac{P(X \in [-5;0] \cap [-5;3])}{P(X \leq 3)} = \frac{P(X \leq 0)}{P(X \leq 3)} = \frac{0{,}5}{0{,}8} = 0{,}625$$

13. Supóngase que X está distribuida uniformemente en [-a,a] con a>0. Cada vez que sea posible, determinar "a" de modo que satisfaga lo siguiente:
   a) $P(X > 1) = \frac{1}{3}$ b) $P(X > 1) = \frac{1}{2}$ c) $P\left(X \leq \frac{1}{2}\right) = 0{,}7$ d) $P\left(X \leq \frac{1}{2}\right) = 0{,}3$
   e) $P(|X| \leq 1) = P(|X| > 1)$

Si una variable aleatoria está distribuida uniformemente en un intervalo [-a;a], el valor de la constante será $\frac{1}{2a}$. Para resolver el punto a) vemos que este parámetro "a" no puede ser menor que "1" porque la probabilidad de que la variable aleatoria sea mayor que "1" sería nula. Luego hay que integrar entre "1" y "a" y esta probabilidad debe dar 1/3.

$$\int_{1}^{a} \frac{1}{2a} dx = \frac{a-1}{2a} = \frac{1}{3}$$

De donde a=3.
Para resolver el punto b) nos encontramos con una planteo similar pero

$$\int_{1}^{a} \frac{1}{2a} dx = \frac{a-1}{2a} = \frac{1}{2}$$

Luego

$$2a - 2 = 2a$$

Llegamos a una contradicción (-2=0), lo que nos dice que no existe un valor de "a" que cumpla la condición pedida.
Para resolver el punto c) planteamos

$$\int_{-a}^{\frac{1}{2}} \frac{1}{2a} dx = \frac{\frac{1}{2}+a}{2a} = 0{,}7$$

De donde

$$2a * 0{,}7 = \frac{1}{2} + a$$

Luego a=1,25. Para el punto d)

$$\int_{-a}^{\frac{1}{2}} \frac{1}{2a} dx = \frac{\frac{1}{2}+a}{2a} = 0{,}3$$

Entonces

$$2a * 0{,}3 = \frac{1}{2} + a$$

Y resulta $a = -1{,}25$. Como “a” tiene que ser positivo por el planteo propuesto, tal coeficiente no existe para la probabilidad propuesta.
Para el punto f) observamos que si a<1 no podría darse la igualdad propuesta. Luego

$$\int_{-1}^{1} \frac{1}{2a} dx = \int_{-a}^{-1} \frac{1}{2a} dx + \int_{1}^{a} \frac{1}{2a} dx$$

Luego

$$\frac{-2}{2a} = \frac{-a+1}{2a} + \frac{1-a}{2a}$$

De donde

$$-2 = 2 - 2a$$

Y $a = 2$ verifica la condición pedida.

14. La función de densidad dada por $f(x) = x/4$ está definida en el intervalo (1, 3). Hallar la función de distribución de $f(x)$. Si $Y = 2X + 3$, hallar las funciones de distribución y de densidad de Y.

Dado que f(x)=x/4 en el intervalo (1;3), $F(x) = \frac{x^2-1}{8}$
Planteamos formalmente

$$G(y) = P(Y \leq y) = P(2X + 3 \leq y) = P\left(X \leq \frac{y-3}{2} = x\right) = F\left(\frac{y-3}{2}\right) = \frac{\left(\frac{y-3}{2}\right)^2 - 1}{8}$$
$$= \frac{y^2 - 6y + 5}{32}$$

Formalmente

$$G(y) = \begin{cases} 1 & 9 \leq y \\ \frac{y^2 - 6y + 5}{32} & 5 < y < 9 \\ 0 & y \leq 5 \end{cases}$$

Para hallar la función de densidad derivamos la de distribución

$$g(y) = \begin{cases} 0 & 9 \leq y \\ \frac{y-3}{16} & 5 < y < 9 \\ 0 & y \leq 5 \end{cases}$$

15. Una corriente eléctrica I, medida en amperes, que fluctúa al azar, se puede considerar como una variable aleatoria distribuida uniformemente en el intervalo real (9; 11). Si esta corriente pasa por un conductor con resistencia de R=2 medida en ohm, encontrar la función de densidad de probabilidad de la potencia, sabiendo que se calcula de acuerdo con la ecuación $P = RI^2$.

Con un planteo similar al anterior y omitiendo unidades, tomando $Y=2X^2$ calculamos $F(x) = \frac{x-9}{2}$ en el intervalo indicado.

$$G(y) = P(Y \leq y) = P(2X^2 \leq y) = P\left(X \leq \sqrt{\frac{y}{2}} = x\right) = \frac{\sqrt{\frac{y}{2}} - 9}{2}$$

Formalmente

$$G(y)=\begin{cases}1 & 242\le y\\ \dfrac{\sqrt{\frac{y}{2}}-9}{2} & 162<y<242\\ 0 & y\le 162\end{cases}$$

Y

$$g(y)=\begin{cases}0 & 242\le y\\ \sqrt{\dfrac{1}{8y}} & 162<y<242\\ 0 & y\le 162\end{cases}$$

Unidad 4: Parametrización de distribuciones

Con el fin de describir las propiedades de las distribuciones es usual resumir tales propiedades en unos pocos parámetros numéricos. Al igual que el análisis de funciones sintetiza propiedades en valores como el rango, extremos, máximos, mínimos, puntos de inflexión, también en el cálculo de probabilidades se seleccionan parámetros como medidas de posición, de variabilidad y de forma que analizaremos a continuación.

El *mínimo* de una distribución es el menor valor que puede adquirir una variable aleatoria, de la misma manera que el *máximo* es el mayor valor posible. Así, al tirar un dado, el mínimo vale 1 y el máximo vale 6. El *rango* de valores posibles, rango de definición o simplemente rango, se obtiene como la diferencia entre el máximo y el mínimo. Al tirar un dado el rango va entre 1 y 6 (seis valores posibles).

Un parámetro usado muy frecuentemente es la *moda* o valor modal, que se interpreta como el valor de variable aleatoria con máximo de probabilidad. En el caso de tratarse de una variable aleatoria continua, su determinación es sencilla y se limita a hallar los máximos de la función de densidad $f_{(x)}$ igualando a cero la derivada primera y mostrando que la derivada segunda es negativa. En el caso de las variables aleatorias discretas, se requiere comparar en términos generales el valor de probabilidad del supuesto máximo $x_{modal} = x_m$ con los inmediatos anterior y posterior x_{m-1} y x_{m+1}, es decir verificar para qué valor de X vale la condición de máximo comparado con los inmediatos laterales $P(X = x_m) \geq P(X = x_{m-1})$ y $P(X = x_m) \geq P(X = x_{m+1})$.

Otro parámetro de uso general se conoce como *mediana* $\tilde{x}$ y es el valor de variable aleatoria que divide la distribución en dos partes iguales. Es decir que $P(X \leq \tilde{x}) = 0{,}5$, o bien que $F_{(\tilde{x})} = 0{,}5$. Esta evaluación no es sencilla de hacer en el caso de variables aleatorias discretas debido a que para un valor particular la función $F_{(x)}$ podría ser menor que 0,5 y para el siguiente podría exceder tal valor, pero en distribuciones continuas basta hallar $\int_{-\infty}^{\tilde{x}} f_{(x)} dx = 0{,}5$, siendo $\tilde{x}$ la incógnita, o bien hallar $\tilde{x}$ tal que $F_{(\tilde{x})} = 0{,}5$.

Sobre la base de la misma idea se definen los *terciles* o valores de variable aleatoria que dividen la distribución en tres partes iguales, es decir que si x_{t1} y x_{t2} son los dos terciles, $F_{(x_{t1})} = \frac{1}{3}$ y $F_{(t_{t2})} = \frac{2}{3}$. También se utilizan *cuartiles* x_{c1}, x_{c2} y x_{c3} tales que $F_{(x_{c1})} = \frac{1}{4}$, $F_{(x_{c2})} = \frac{2}{4}$ y $F_{(x_{c3})} = \frac{3}{4}$. Puede verse que el segundo cuartil coincide con la mediana. También se utilizan *quintiles* x_{q1} hasta x_{q4}, *deciles*, x_{d1} a x_{d9} y *percentiles* x_{p1} a x_{p99}. Aquí el percentil cincuenta coincide con la mediana. En la práctica se usan como medidas rústicas que aproximan valores de probabilidad cuando no se puede pedir mucha precisión a la información disponible. Por ejemplo los terciles separan el tercio central del tercio más pequeño y el tercio más grande, los quintiles hacen algo parecido con un poco más de resolución separando el veinte por ciento más alejado en ambos extremos del veinte por ciento central y dos medidas de apartamiento de la centralización no tan extremas en los quintos segundo y cuarto. Los deciles subdividen medidas de probabilidad con resolución de un diez por ciento y los percentiles con un uno por ciento, aunque para poder llegar a tal nivel de resolución se requieren mucha información y se aplica a situaciones en que la aproximación a la distribución de probabilidad es casi totalmente empírica. Veremos que los cuartiles tienen algunas aplicaciones como medidas de variabilidad y de forma.

La moda y la mediana son medidas de posición, de localización o de centralización, en el sentido que definen la ubicación de la distribución dentro del eje real. El rango $Rg = x_{máx} - x_{mín}$, el rango intercuartílico $x_{c3} - x_{c1}$ y otras medidas similares caracterizan la dispersión de la distribución, el primero delimitando el rango máximo de dispersión de la variable en cuestión y el segundo el rango donde se distribuye el 50% de la probabilidad. Pero también hay medidas de forma, como la posición relativa entre la moda y la mediana o los rangos intercuartílicos laterales $x_{c3} - x_{c2}$ y $x_{c2} - x_{c1}$ que analizaremos luego.

Esperanza

El concepto de esperanza es uno de los más útiles aunque su interpretación resulta un poco más oscura que las de otros parámetros ya mencionados. Inclusive el nombre *esperanza* induce a error debido a que no se trata de ningún valor que se “espere” obtener como resultado de un experimento aleatorio.

La esperanza es un parámetro de posición definido como una manera de determinar el punto de equilibrio de una distribución puntual o de una función de densidad. En tal sentido es una medida de posición o de localización central de la distribución. Se la interpreta como la ubicación en la cual debería encontrarse toda la distribución de probabilidad asociada a un experimento si ésta se concentrarse en un único punto. En este sentido nos permite expresarnos en un lenguaje determinista como si toda la distribución estuviese concentrada en el valor esperado.

Matemáticamente se define

$$E_{(x)} = \sum_{i=1}^{n} x_i P_{(X=x_i)}$$

si la distribución es discreta y

$$E_{(x)} = \int_{-\infty}^{+\infty} x f_{(x)} dx$$

si la distribución es continua.

Puede verse que la definición es similar en forma a la que se da del momento de una fuerza o al centro de masa o al centro de gravedad. Basta reemplazar x por la distancia a un punto y $P_{(X=x)}$ por la fuerza aplicada para obtener la definición del momento, o a $f_{(x)}dx$ por un diferencial de masa o de fuerza peso para obtener el centro de masa o el de gravedad. Es frecuente que se denomine a la esperanza “momento no centrado de primer orden”.

Como ejemplo de aplicación para una variable aleatoria discreta, calculemos la esperanza del experimento consistente en tirar un dado. Los resultados posibles son los conocidos {1, 2, 3, 4, 5, 6} con probabilidades $P(X = i) = 1/6$ en todos los casos. Por definición

$$E(X) = \sum_{i=1}^{6} i * P(X = i) = 1 * \frac{1}{6} + 2 * \frac{1}{6} + \cdots + 6 * \frac{1}{6} = \frac{21}{6} = 3{,}5$$

Notemos que $\mu = E(X) = 3{,}5$ no es un resultado posible pero define un punto de equilibrio que deja a la misma distancia los tres primeros resultados (1;2;3) y los tres últimos resultados (4;5;6).

En un ejemplo sobre una distribución continua, sea $f_{(x)} = 3x^2/8$ con $0 \leq x \leq 2$, primero verificamos que es una función de densidad. Es claro que es no negativa y

$$\int_{-\infty}^{+\infty} \frac{3x^2}{8} dx = \int_{-\infty}^{0} 0.dx + \int_{0}^{2} \frac{3x^2}{8} dx + \int_{2}^{+\infty} 0.dx = \frac{x^2}{8}\Big|_0^2 = 1$$

La esperanza resulta

$$E_{(x)} = \int_{-\infty}^{+\infty} x f_{(x)} dx$$

$$E_{(x)} = \int_{-\infty}^{0} x.0.dx + \int_{0}^{2} x\frac{3x^3}{8} dx + \int_{2}^{+\infty} x.0.dx = 0 + \frac{3}{8}\frac{2^4}{4} + 0 = \frac{3}{2} \cong 1{,}5$$

Bastará graficar el arco de parábola, con mayor acumulación de probabilidad hacia la derecha, para visualizar que es un punto de equilibrio.

Varianza y Dispersión

En términos más generales, puede definirse la esperanza de cualquier función de variable aleatoria en la forma

$$E_{(H(x))} = \int_{-\infty}^{+\infty} H_{(x)} f_{(x)} dx$$

y

$$E_{(H(x))} = \sum_{i=1}^{n} H_{(x_i)} P_{(X=x_i)}$$

para variables aleatorias continuas y discretas respectivamente.

La varianza es un caso particular de una función de una variable aleatoria y a la vez una medida de variabilidad. Es la esperanza de la diferencia cuadrática con respecto a la esperanza de la distribución. Es decir,

$$V_{(x)} = E\left[\left(x - E_{(x)}\right)^2\right]$$

Para una distribución discreta queda

$$V_{(x)} = \sum_{i=1}^{n} \left(x_i - E_{(x)}\right)^2 P_{(X=x_i)}$$

Y para una distribución continua

$$V_{(x)} = \int_{-\infty}^{+\infty} \left(x - E_{(x)}\right)^2 f_{(x)} dx$$

También se la denomina “momento de segundo orden centrado en la esperanza”. Es usual que en estadística se exprese la varianza como σ^2 (sigma cuadrado) y a la esperanza como μ (mu).

Esta definición conduce a ver la varianza como una ponderación, por medio de la probabilidad, de las distancias cuadráticas de la variable aleatoria con respecto a la esperanza de la distribución. En tal sentido será mayor en la medida que la distribución sea dispersa, es decir, que existan muchos valores de la variable aleatoria alejados de la esperanza y con alta probabilidad de ocurrencia. Sobre la base de la varianza se define la dispersión, también llamada dispersión estándar de una distribución, como la raíz cuadrada de la varianza. Usualmente se la expresa con el símbolo griego σ (sigma).

$$\sigma = \sqrt{V_{(x)}} = \sqrt{\sigma^2}$$

El objetivo de calcular la dispersión estándar es recuperar las dimensiones (unidades) de la variable aleatoria y así disponer de una medida de dispersión comparable con el rango, esperanza y otros parámetros, expresados en las unidades propias de la variable aleatoria. Así, si la variable aleatoria es una distancia, se medirá en metros, pero su varianza quedará expresada en metros cuadrados. Sin embargo la varianza es una medida de variabilidad de distancias, no una medida de superficie. De allí que la raíz cuadrada de la varianza hace comparables la dispersión resultante con los otros parámetros.

Calculemos la varianza del experimento “tirar un dado”.

$$\sigma^2 = V(X) = E((X - \mu)^2) = \sum_{i=1}^{6} (i - 3{,}5)^2 P(X = i) =$$

$$(1-3{,}5)^2 * \frac{1}{6} + (2-3{,}5)^2 * \frac{1}{6} + \cdots + (6-3{,}5)^2 * \frac{1}{6} =$$
$$6{,}25 * \frac{1}{6} + 2{,}25 * \frac{1}{6} + 0{,}25 * \frac{1}{6} + 0{,}25 * \frac{1}{6} + 2{,}25 * \frac{1}{6} + 6{,}25 * \frac{1}{6} = 2{,}91\hat{6}$$

A partir de este valor de la varianza calculamos la dispersión estándar como raíz cuadrada

$$\sigma = \sqrt{\sigma^2} = \sqrt{2{,}916\ldots} \cong 1{,}708$$

Si calculamos

$$\mu - \sigma \cong 3{,}5 - 1{,}7 = 1{,}8$$
$$\mu + \sigma = 3{,}5 + 1{,}7 = 5{,}2$$

Vemos que dentro del intervalo definido por estos dos límites se encuentra los valores 2, 3, 4 y 5 de nuestra variable aleatoria, dos tercios de los seis casos posibles.

Calculemos la varianza de la distribución propuesta previamente como $f_{(x)} = x^2/3$ con $0 \leq x \leq 2$.

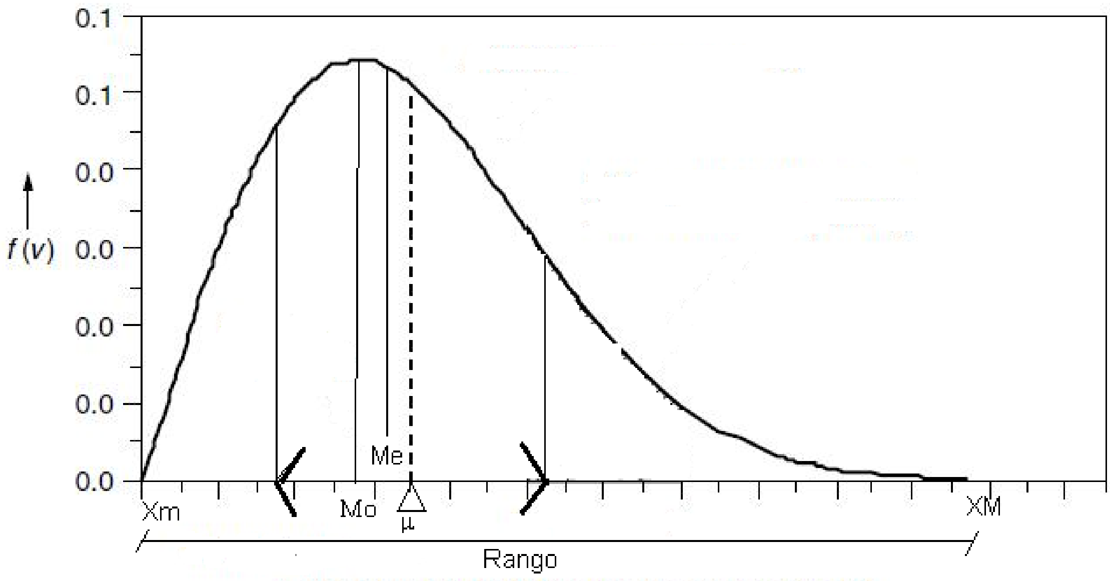
$$V_{(x)} = \int_{-\infty}^{0} \left(x - \frac{3}{2}\right)^2 .0.dx + \int_{0}^{2} \left(x - \frac{3}{2}\right)^2 .\frac{3x^2}{8}.dx + \int_{2}^{+\infty} \left(x - \frac{3}{2}\right)^2 .0.dx = 0{,}15$$

Dado que $\sigma^2 = 0{,}15$, resulta en el ejemplo $\sigma = 0{,}39$. Puede verse que en este caso si a la esperanza, de valor 1,5, se le suma y resta el valor de la dispersión se obtienen las cotas 1,50..-0,39=1,11.. y 1,50+0,39=1,89... Si calculamos la probabilidad de obtener resultados en el intervalo [1,11..; 1,89..] resulta

$$P(1{,}11.. < X \leq 1{,}89..) = F_{(1{,}89..)} - F_{(1{,}11..)} = \frac{(1{,}89..)^3}{8} - \frac{(1{,}11..)^3}{8} = \frac{(2{,}532 - 0{,}513)}{8}$$
$$\cong 0{,}673$$

Es decir que hay casi un 67,3% de probabilidad de que se obtenga un valor entre la esperanza y los límites definidos por $\mu - \sigma$ y $\mu + \sigma$. Así la varianza es una medida de dispersión, muy útil a los fines de cálculo pero de difícil interpretación. La dispersión, en cambio, es de más sencilla interpretación y comparable con el rango de variabilidad en términos de las dimensiones de la variable aleatoria. Los límites $\mu - \sigma$ y $\mu + \sigma$ también son de sencilla interpretación y pueden ser graficados en la misma escala que la variable aleatoria. Estos límites se encuentran dentro del rango de la variable aleatoria y la probabilidad de obtener resultados dentro de tal intervalo es tanto mayor cuanto más concentrada, unimodal y simétrica es la distribución. En distribuciones típicas moderadamente simétricas y unimodales usualmente se encuentra una probabilidad del orden del 60% a 70% entre los límites $\mu - \sigma$ y $\mu + \sigma$.

En el siguiente gráfico de una distribución aproximada de módulo de velocidad del viento, ligeramente asimétrica en sentido positivo, ubicamos la esperanza, $\mu \pm \sigma$ por medio de paréntesis, la mediana, la moda, los extremos y el rango.



A windspeed probability density function (p.d.f).

Una forma a veces práctica de calcular la varianza nos lleva a un pequeño teorema, que demostraremos sólo para una distribución continua. Desarrollamos la expresión

$$V_{(x)} = \int_{-\infty}^{+\infty} \left(x - E_{(x)}\right)^2 f_{(x)} dx = \int_{-\infty}^{+\infty} x^2 f_{(x)} dx - \int_{-\infty}^{+\infty} 2E_{(x)} x f_{(x)} dx + \int_{-\infty}^{+\infty} E_{(x)}^2 f_{(x)} dx$$

Extraemos de las integrales las constantes numéricas

$$\int_{-\infty}^{+\infty} x^2 f_{(x)} dx - 2E_{(x)} \int_{-\infty}^{+\infty} x f_{(x)} dx + E_{(x)}^2 \int_{-\infty}^{+\infty} f_{(x)} dx =$$

La primera integral la escribimos como $E_{(x^2)}$, la segunda integral es la esperanza $E_{(x)}$ y la tercera integral vale 1 por ser una función de densidad, de modo que

$$= E_{(x^2)} - 2E_{(x)}E_{(x)} + E_{(x)}^2 .1 = E_{(x^2)} - 2E_{(x)}^2 + E_{(x)}^2 = E_{(x^2)} - E_{(x)}^2$$

Al parámetro $E_{(x^2)}$ suele llamarse el “momento de segundo orden de la distribución”, concepto que generalizaremos más adelante.

También demostraremos que la varianza es el momento centrado de segundo orden mínimo cuando el parámetro de centralización es la esperanza. Nos limitaremos a demostrarlo para una distribución continua. Sea c un parámetro de centralización genérico para un momento centrado de segundo orden, que notaremos $\mu_{2(c)}$

$$\mu_{2(c)} = \int_{-\infty}^{+\infty} (x - c)^2 f_{(x)} dx$$

Si pensamos que este momento depende del parámetro de centralización, para determinar un mínimo lo derivamos con respecto a c, igualamos la derivada a cero y verificamos que la derivada segunda sea positiva.

$$\frac{d\mu_{2(c)}}{dc} = -\int_{-\infty}^{+\infty} 2(x - c) f_{(x)} dx = 0 = -2 \int_{-\infty}^{+\infty} x f_{(x)} dx + 2 \int_{-\infty}^{+\infty} c f_{(x)} dx$$

De modo que, si omitimos el coeficiente 2 y el signo,

$$\int_{-\infty}^{+\infty} x f_{(x)} dx = E_{(x)} = c \int_{-\infty}^{+\infty} f_{(x)} dx = c.1 = c$$

Esto prueba que $c = E_{(x)}$ es un extremo y, si derivamos nuevamente obtenemos

$$\frac{d^2 \mu_{2(c)}}{dc^2} = 2 \int_{-\infty}^{+\infty} f_{(x)} dx = 2.1 = 2 > 0$$

Lo que prueba que la esperanza, como parámetro de centralización, define un mínimo para el momento centrado de segundo orden o que la varianza es el mínimo momento centrado de segundo orden.

Otra medida de dispersión se conoce como desviación media. Se define como

$$DM = \int_{-\infty}^{+\infty} |x - \tilde{x}| f_{(x)} dx \quad o \quad DM = \sum_{i=0}^{n} |x_i - \tilde{x}| P_{(X=x_i)}$$

según la distribución sea continua o discreta. En estas expresiones $\tilde{x}$ es la mediana de la distribución y las barras corresponden al módulo de la diferencia entre la variable aleatoria y la mediana.

Una medida de dispersión de uso frecuente pero que debe interpretarse con precaución cuando la esperanza es próxima a cero es la variabilidad relativa, que es adimensional y puede expresarse en porcentaje

$$VR = \frac{\sigma}{\mu}$$

Otra medida equivalente pero sustentada sobre la mediana y la desviación media es

$$VR^* = \frac{DM}{\tilde{x}}$$

Aun medidas más simples pero no siempre fáciles de utilizar en la práctica y limitadas al manejo de variables con un solo signo son

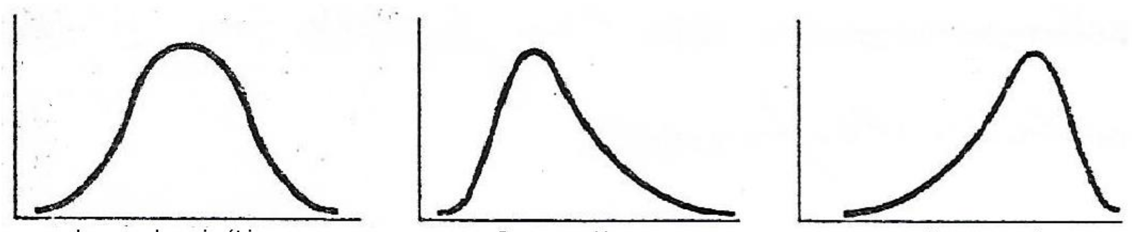
$$Q_1 = \frac{Máximo}{Mínimo}$$

y

$$Q_2 = \frac{Rango}{\mu}$$

Medidas de forma

Se interpreta como medidas de forma a ciertos parámetros que caracterizan el aspecto de la distribución. Esto resulta complejo y muy variado. La primera consideración acerca de la forma se relaciona con el apartamiento de la simetría. Se dice que una distribución es “insesgada” o “simétrica” si presenta simetría con respecto a un eje central que coincide con la esperanza y con la mediana, y usualmente con la moda. Si presenta un ascenso rápido y un lento decrecimiento se dice que tiene “sesgo positivo” y, si el crecimiento es lento con un decrecimiento rápido, el “sesgo” es “negativo”, según se ve en las figuras que siguen.



Hay varias maneras de asociar un parámetro a estas diferencias de forma. Uno de los que más se usa se define a partir de la noción de momento centrado. Extendemos la potencia de las diferencias con respecto a la esperanza al tercer orden de modo que se recupera el signo de la desviación y se potencia la medida de distancia. Se define así como "momento centrado en la esperanza de tercer orden" a

$$\mu_3 = \int_{-\infty}^{+\infty} \left(x - E_{(x)}\right)^3 f_{(x)} dx \qquad o \qquad \mu_3 = \sum_{i=1}^{n} \left(x_i - E_{(x)}\right)^3 P_{(X=x_i)}$$

según sea continua o discreta la variable aleatoria. Podría sacarse la raíz cúbica para recuperar la unidad comparable con la esperanza y la desviación, pero en su lugar se define un parámetro adimensional que se conoce con el nombre de "asimetría" o "sesgo". Para ello se divide el momento de tercer orden centrado en la esperanza por el cubo del desvío estándar, también centrado en la esperanza. Según sea continua o discreta tenemos

$$\gamma_{1(x)} = \frac{\int_{-\infty}^{+\infty} \left(x - E_{(x)}\right)^3 f_{(x)} dx}{\sigma^3} \quad o \quad \gamma_{1(x)} = \frac{\sum_{i=1}^{n} \left(x_i - E_{(x)}\right)^3 P_{(X=x_i)}}{\sigma^3}$$

Este parámetro tiene un valor nulo si la distribución es simétrica. Si el gráfico de la distribución se "estira" hacia la derecha, tiene asimetría positiva, pero si se "estira" hacia la izquierda, tiene asimetría negativa.

En el ejemplo desarrollado más arriba vale

$$\gamma_{1(x)} = \frac{\int_0^2 \left(x - \frac{3}{2}\right)^3 \frac{3x^2}{8} dx}{0{,}39^3} = \frac{-0{,}04875}{0{,}39^3} = -0{,}82$$

Se obtiene en este caso una asimetría claramente negativa.

Una medida de forma menos objetiva pero sencilla está dada por la ubicación relativa entre la esperanza, la mediana y la moda cuando es unimodal. Si la distribución tiene sesgo positivo el orden típico es moda-mediana-esperanza, y, si tiene sesgo negativo es esperanza-mediana-moda. En la figura que vimos más arriba relativa a la distribución de velocidades del viento mostramos este orden en una curva que representa una asimetría claramente positiva.

También la posición relativa de los cuartiles es una medida de forma, así, si la distancia entre el primero y segundo cuartil es menor que entre el segundo y tercero, la distribución tiene asimetría positiva, mientras que lo contrario ocurre si tiene asimetría negativa, y si son iguales será simétrica, y en tal caso también coinciden la mediana y la esperanza, no necesariamente la moda a menos que sea unimodal.

$$\begin{gathered} Si\ x_{c2} - x_{c1} < x_{c3} - x_{c2} \ \rightarrow A > 0 \\ Si\ x_{c2} - x_{c1} = x_{c3} - x_{c2} \ \rightarrow A = 0 \\ Si\ x_{c2} - x_{c1} > x_{c3} - x_{c2} \ \rightarrow A < 0 \end{gathered}$$

También se puede expresar como cociente. Si

$$\frac{x_{c3} - x_{c2}}{x_{c2} - x_{c1}} > 1 \rightarrow \gamma_1 > 0$$

$$\frac{x_{c3} - x_{c2}}{x_{c2} - x_{c1}} = 1 \rightarrow \gamma_1 = 0$$

$$\frac{x_{c3} - x_{c2}}{x_{c2} - x_{c1}} < 1 \rightarrow \gamma_1 < 0$$

Curtosis o apuntalamiento

Hemos notado quizá que el símbolo usado para la asimetría fue γ_1. No es unánime la notación pero es muy frecuente. Ésta sugiere que existe un γ_2 al menos. Este parámetro se define de acuerdo con la siguiente relación.

$$\gamma_{2(x)} = \frac{E\left[\left(x - E_{(x)}\right)\right]^4}{\sigma^4} - 3$$

El significado de este parámetro se estudiará más adelante porque se define en relación con una curva descriptiva de lo que se conoce como “distribución normal” o “de Gauss”, o “campana de Gauss”. Esta distribución, de muy amplio uso en el cálculo de probabilidades y de la estadística, tiene una curtosis de “3”. El primer término sigue la idea de la esperanza del momento de cuarto orden centrado en la esperanza con respecto a la cuarta potencia de la desviación estándar. Al restar el valor correspondiente a la distribución normal, resulta ser un parámetro de comparación con esta distribución según sea mayor, igual o menor que cero, y lo estudiaremos después de definirla.

Esperanza y varianza de transformaciones lineales de variables aleatorias

Una transformación lineal de una variable aleatoria responde a un cambio de variable de la forma $Y = aX + b$, resulta

$$E_{(Y)} = E_{(aX+b)} = \int_{-\infty}^{+\infty} (ax + b) f_{(x)} dx = a \int_{-\infty}^{+\infty} x f_{(x)} dx + b \int_{-\infty}^{+\infty} f_{(x)} dx = aE_{(X)} + b$$

por definición de esperanza y propiedades de la función de densidad.

Para la varianza

$$\begin{aligned} V_{(Y)} = V_{(aX+b)} &= E\left[\left((aX + b) - E(aX + b)\right)^2\right] = E\left[\left(aX + b - aE_{(X)} - b\right)^2\right] \\ &= E\left[a^2\left(X - E_{(X)}\right)^2\right] = a^2 E\left[\left(X - E_{(X)}\right)^2\right] = a^2 V_{(X)} \end{aligned}$$

Es decir que tenemos dos relaciones generales que relacionan la esperanza y la varianza de una transformación lineal de variables aleatorias.

$$E_{(aX+b)} = aE_{(X)} + b \qquad y \qquad V_{(aX+b)} = a^2 V_{(X)}$$

Puede verse así que si $a = \frac{1}{\sigma}$ y $b = -\frac{\mu}{\sigma}$, siendo μ la esperanza de X y σ^2 la varianza de X, resulta $Y = aX + b = \frac{X-\mu}{\sigma}$. A esta transformación lineal se la conoce como “estandarización” de una variable aleatoria, es decir que *Y* es la variable aleatoria *X* estandarizada de manera tal que $E_{(Y)} = 0$ y $V_{(Y)} = 1$.

Parámetros de la distribución binomial

La esperanza de una distribución binomial se obtiene como $E_{(x)} = np$, donde n es el número de repeticiones y p la probabilidad de éxito. Trataremos de demostrar esta relación. Si $X \sim Bi(n,p)$ contabiliza el número de éxitos en n repeticiones independientes de un experimento binomial

$$\mu = E(X) = \sum_{i=0}^{n} i \binom{n}{i} p^i (1-p)^{n-i} = \sum_{i=0}^{n} i \frac{n!}{i!\,(n-i)!} p^i (1-p)^{n-i} =$$

Dado que el primer término es nulo si $i = 0$,

$$= \sum_{i=1}^{n} i \frac{n!}{i!\,(n-i)!} p^i (1-p)^{n-i} = \sum_{i=1}^{n} in \frac{(n-1)!}{i(i-1)!\,\big(n-1-(i-1)\big)!} pp^{i-1} (1-p)^{n-1-(i-1)} =$$

y llamando $j = i - 1, m = n - 1$

$$= \sum_{j=0}^{m} n \frac{m!}{j!\,(m-j)!} pp^j (1-p)^{m-j} = np \sum_{j=0}^{m} \binom{m}{j} p^j (1-p)^{m-j} = np * 1 = np$$

Luego

$$\mu = E(X) = np$$

La varianza de una distribución binomial se obtiene de sus parámetros por medio de $V_{(x)} = np(1-p)$. La demostración es un poco más extensa pero la veremos más delante de una manera simple

$$\sigma^2 = V(X) = np(1-p)$$

Se puede demostrar que el momento centrado de tercer orden está dado por

$$\mu_3 = 2np(1-p)\left(\frac{1}{2} - p\right)$$

De modo que la asimetría resulta

$$\gamma_1 = 2 \frac{\frac{1}{2} - p}{\sqrt{np(1-p)}}$$

Y el momento centrado de cuarto orden queda

$$\mu_4 = np(1-p)\big(1 + 3p(1-p)(n-2)\big)$$

Por lo tanto la curtosis

$$\gamma_2 = \frac{1 - 6p(1-p)}{np(1-p)}$$

Momentos

Hemos mencionado anteriormente el término “momento”. Tiene la finalidad de definir parámetros en forma muy general. Se define como “momento no centrado de orden ν”, según la variable aleatoria sea discreta o continua, a

$$\alpha_\nu = E(X^\nu) = \sum_{i=0}^{n} x_i^\nu P(X = x_i) \qquad o \qquad \alpha_\nu = E(X^\nu) = \int_{-\infty}^{+\infty} x^\nu f(x)dx$$

A los fines de ejemplificar nos limitaremos a una variable aleatoria continua. El momento de orden “cero” vale “uno”.

$$\alpha_0 = \int_{-\infty}^{+\infty} x^0 f(x)dx = \int_{-\infty}^{+\infty} f(x)dx = 1$$

El momento de primer orden es la esperanza

$$\alpha_1 = \int_{-\infty}^{+\infty} x^1 f(x)dx = \int_{-\infty}^{+\infty} xf(x)dx = \mu$$

Los momentos de orden superior no tienen una denominación pero serán utilizados.
Por otra parte se llama “momento centrado en ‘c’ de orden ν” a

$$\mu_{c\nu} = E((X-c)^\nu) = \sum_{i=0}^{n} (x_i - c)^\nu P(X = x_i) \quad c \quad \mu_{c\nu} = \int_{-\infty}^{+\infty} (x-c)^\nu f(x)dx$$

Si $c = \mu = E(X)$ se lo llama “momento centrado en la esperanza de orden ν”. Ya hemos demostrado que la esperanza es define un mínimo para los momentos centrados. Cuando el parámetro de centralización sea la esperanza, se omitirá en la notación. Si $\nu = 2$ tenemos la definición de “varianza”.

$$\mu_2 = \sigma^2 = E((X-\mu)^2) = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx = \alpha_2 - \mu^2$$

La última igualdad la habíamos demostrado en forma general. Si tenemos presente que la expresión $\mu_{c2} = \int_{-\infty}^{+\infty}(x-c)^2 f(x)dx$ define una medida de norma y, por lo tanto, de distancia a “c” ponderada por la probabilidad, que sea mínimo cuando $c = \mu$, nos dice que la esperanza es un parámetro que se encuentra a una distancia mínima de todos los valores posibles de la distribución de probabilidad.

De la misma manera que probamos una relación entre la varianza y los momentos, también lo podemos con los momentos centrados en la esperanza de tercero y cuarto orden

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

Y

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

Momentos y parámetros de una distribución uniforme

Si calculamos los momentos de una distribución uniforme u[a,b] genérica tenemos

$$\alpha_\nu = \frac{1}{b-a}\int_a^b x^\nu f(x)dx = \frac{1}{\nu+1}\frac{b^{\nu+1} - a^{\nu+1}}{b-a}$$

Tendremos

$$\alpha_1 = \mu = \frac{1}{2}\frac{b^2 - a^2}{b - a} = \frac{1}{2}\frac{(b + a)(b - a)}{b - a} = \frac{a + b}{2}$$

Luego

$$\mu_2 = \sigma^2 = \alpha_2 - \mu^2 = \frac{1}{2 + 1}\frac{b^{2+1} - a^{2+1}}{b - a} - \frac{(a + b)^2}{2^2} = \frac{(b - a)^2}{12}$$

$$\mu_3 = 0$$

Y

$$\mu_4 = \frac{(b - a)^4}{80}$$

Probabilidades y Estadística: Ejercicios Unidad Nº 4

Parametrización de Distribuciones

1. Se tiran cuatro monedas y contabiliza el número de caras obtenidas. Hallar la esperanza y varianza de la variable aleatoria “número de caras”. Se tira un dado tres veces y considera éxito si se obtiene un seis. Hallar la esperanza y varianza del número de éxitos y verificar el cálculo de la esperanza y varianza de una binomial como *np* y *np(1-p)*.

2. Un lote de 10 motores eléctricos debe ser rechazado totalmente o bien vendido, según el resultado del siguiente proceso: se escogen al azar dos motores y se inspeccionan. Si uno o más son defectuosos, el lote es rechazado; de otro modo es aceptado. Supóngase que cada uno de los motores cuesta $75 y se vende por $100; si el lote contiene un motor defectuoso, ¿cuál es la utilidad esperada por el fabricante? Hallar la esperanza del número de motores defectuosos encontrados. Hallar la varianza y desvío estándar de la utilidad y del número de motores defectuosos. Graficar las distribuciones y los parámetros esperanza y desvío. Hallar la asimetría del número de motores defectuosos y de la utilidad.

3. Un cable se fabrica con un diámetro variable en un rango entre 6 mm y 8 mm que responde a una función de densidad polinómica de la forma

$$f_{(x)} = a(x-4)$$

a) Hallar el valor de la constante “a”.
b) Graficar la densidad y la distribución.
c) Calcular el valor esperado μ de la variable aleatoria “diámetro”.
d) Calcular la varianza y la desviación estándar σ de dicha variable aleatoria.
e) Calcular la asimetría de la variable aleatoria.
f) Calcular la probabilidad de que el diámetro sea mayor que 7.
g) Calcular la probabilidad de que el diámetro pertenezca al intervalo [6,5;7,5].
h) Indicar en los gráficos la esperanza, desvío y probabilidades calculadas.
i) Hallar la probabilidad asociada al intervalo [μ-σ; μ+σ]
j) Hallar la mediana y los cuartiles de la distribución.

4. Supóngase que X está distribuida uniformemente en [0,a]. Encontrar la esperanza μ y varianza σ^2 de X. Hallar la probabilidad de que la variable aleatoria se encuentre en el intervalo [μ-σ; μ+σ]

5. Supóngase que un instrumento electrónico tiene una duración X (en unidades de 1000 horas) que se considera como una variable aleatoria continua con la siguiente función de densidad:

$$f_{(x)} = e^{-x} \quad si\ x > 0$$

Suponiendo que el costo de fabricación de tal artículo es $2.00. El fabricante vende el artículo por $5.00, pero garantiza un reembolso total si $X \leq 0{,}9$. ¿Cuál es la utilidad esperada del fabricante por artículo? Hallar la varianza de la utilidad, el desvío estándar y la asimetría. Graficar la distribución de la duración y de la utilidad.

6. Cierta aleación se forma al combinar la mezcla fundida de dos metales. La aleación que resulta contiene cierto porcentaje de plomo *X*, que puede considerarse como una variable aleatoria. Supongamos que *X* tiene la siguiente función de densidad:

$$f_{(x)} = \frac{3}{5}10^{-5}x(100-x) \quad para\ 0 \leq x \leq 100$$

Suponemos que P es la utilidad neta obtenida al vender esta aleación (por kg) y está dada por la siguiente función del porcentaje del contenido de plomo $P = C_1 + C_2X$, en donde $C_1 = 100$ y $C_2 = 2$. Graficar la densidad, hallar y graficar la distribución del contenido de plomo, calcular su esperanza, varianza, desvío, mediana, cuartiles, moda y asimetría. Calcular la utilidad esperada y varianza (por kg).

7. Suponiendo que X tiene una función de densidad:

$$f_{(x)} = \frac{x-4}{6} \quad para\; 6 < x < 8$$

Sea $W = \frac{1}{3}X$

a) Calcular $E_{(W)}$ usando la función de densidad de X.

b) Calcular $E_{(W)}$ usando la función de densidad de W.

8. Supóngase que X es una variable aleatoria para la cual $E_{(X)} = 10$ y $V_{(X)} = 25$. Si se define la variable aleatoria $Y = aX + b$, hallar la relación entre la esperanza y varianza de Y y de X. ¿Para qué valores positivos de a y b tiene $Y = aX + b$ esperanza 0 y varianza 1? (Se dice que la variable Y está estandarizada).

Parametrización de Distribuciones: Resolución de ejercicios

1. Se tiran cuatro monedas y contabiliza el número de caras obtenidas. Hallar la esperanza y varianza de la variable aleatoria "número de caras". Se tira un dado tres veces y considera éxito si se obtiene un seis. Hallar la esperanza y varianza del número de éxitos y verificar el cálculo de la esperanza y varianza de una binomial como *np* y *np(1-p)*.

Primero definimos el rango de la variable aleatoria $X = \{0; 1; 2; 3; 4\}$, luego calculamos la probabilidad para cada valor "k" de "X" a partir de una distribución binomial dado que la moneda es equilibrada $X \sim Bi\left(4; \frac{1}{2}\right)$

$$P(X = k) = \binom{4}{k}\left(\frac{1}{2}\right)^{k}\left(1 - \frac{1}{2}\right)^{4-k} = \binom{4}{k}\left(\frac{1}{2}\right)^{4} = \frac{1}{16}\binom{4}{k}$$

Obtenemos

k	0	1	2	3	4
P(X=k)	1/16=0,0625	4/16=0,25	6/16=0,375	4/16=0,25	1/16=0,0625

A partir de esta tabla construimos el gráfico de la distribución y calculamos

$$E(X) = \sum_{k=0}^{4} kP(X = k) = 0 * \frac{1}{16} + 1 * \frac{4}{16} + 2 * \frac{6}{16} + 3 * \frac{4}{16} + 4 * \frac{1}{16} = \frac{4 + 12 + 12 + 4}{16}$$
$$= \frac{32}{16} = 2$$

Si utilizamos la expresión de la esperanza $E(X) = np = 4 * \frac{1}{2} = 2$

La varianza la obtenemos de

$$V(X) = \sum_{k=0}^{4} \left(k - E(X)\right)^{2} P(X = k) =$$

Desarrollando numéricamente

$$= (0-2)^2 * \frac{1}{16} + (1-2)^2 * \frac{4}{16} + (2-2)^2 * \frac{6}{16} + (3-2)^2 * \frac{4}{16} + (4-2)^2 * \frac{1}{16} =$$

$$= 4 * \frac{1}{16} + 1 * \frac{4}{16} + 0 * \frac{6}{16} + 1 * \frac{4}{16} + 4 * \frac{1}{16} = \frac{16}{16} = 1$$

De aquí resulta que

$$\sigma = \sqrt{V(X)} = \sqrt{1} = 1$$

Podemos aplicar un breve teorema que utiliza el momento de segundo orden como una forma simplificada de calcular la varianza

$$V(X) = E(X^2) - E^2(X) = \mu_2 - \mu_1^2$$

Necesitamos el valor de $\mu_2 = E(X^2)$

$$E(X^2) = \sum_{k=0}^{4} k^2 P(X = k) = 0^2 * \frac{1}{16} + 1^2 * \frac{4}{16} + 2^2 * \frac{6}{16} + 3^2 * \frac{4}{16} + 4^2 * \frac{1}{16} =$$

$$= \frac{4}{16} + \frac{24}{16} + \frac{36}{16} + \frac{16}{16} = \frac{80}{16} = 5$$

Luego

$$V(X) = E(X^2) - E^2(X) = 5 - 2^2 = 1$$

También podemos aplicar la expresión general de la varianza de una binomial

$$V(X) = np(1-p) = 4 * \frac{1}{2} * \left(1 - \frac{1}{2}\right) = 1$$

Como extensión del ejercicio podemos calcular la asimetría y la curtosis

$$\gamma_1 = 2\frac{\frac{1}{2} - p}{\sqrt{np(1-p)}} = 2\frac{\frac{1}{2} - \frac{1}{2}}{\sqrt{4\frac{1}{2}\left(1 - \frac{1}{2}\right)}} = 0$$

$$\gamma_2 = \frac{1 - 6p(1-p)}{np(1-p)} = \frac{1 - 6\frac{1}{2}\left(1 - \frac{1}{2}\right)}{4\frac{1}{2}\left(1 - \frac{1}{2}\right)} = \frac{1 - \frac{3}{2}}{1} = -\frac{1}{2} = -0{,}5$$

Para la segunda parte del ejercicio sólo resolvemos la tabla. Primero definimos el rango de la variable aleatoria $X = \{0; 1; 2; 3\}$, luego calculamos la probabilidad para cada valor “k” de “X” a partir de una distribución binomial dado que el dado es equilibrado $X \sim Bi\left(3; \frac{1}{6}\right)$

$$P(X = k) = \binom{3}{k}\left(\frac{1}{6}\right)^k \left(1 - \frac{1}{6}\right)^{3-k} = \binom{3}{k}\left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{3-k} = \left(\frac{1}{6}\right)^3 5^{3-k}\binom{3}{k}$$

Obtenemos

K	0	1	2	3
P(X=k)	$(5/6)^3$=0,579	$75/6^3$=0,347	$15/6^3$=0,069	$1/6^3$=0,005

Podemos desarrollar todos los cálculos como práctica, lo que se recomiendo, y simplemente utilizar las expresiones válidas para la binomial para verificar los resultados.

$$E(X) = np = 3 * \frac{1}{6} = 0{,}5$$

$$V(X) = np(1-p) = 3 * \frac{1}{6} * \left(1 - \frac{1}{6}\right) = \frac{15}{36} = 0{,}417$$

Si hemos realizado el gráfico, veremos que es fuertemente asimétrica.

$$\gamma_1 = 2\frac{\frac{1}{2} - \frac{1}{6}}{\sqrt{3\frac{1}{6}\left(1 - \frac{1}{6}\right)}} = 2\frac{\frac{2}{6}}{\sqrt{\frac{15}{36}}} = \frac{4}{\sqrt{15}} = \sqrt{\frac{16}{15}} = 1{,}0328$$

$$\gamma_2 = \frac{1 - 6p(1-p)}{np(1-p)} = \frac{1 - 6\frac{1}{6}\left(1 - \frac{1}{6}\right)}{3\frac{1}{6}\left(1 - \frac{1}{6}\right)} = \frac{1 - \frac{5}{6}}{\frac{15}{36}} = \frac{6}{15} = 0{,}4$$

2. Un lote de 10 motores eléctricos debe ser rechazado totalmente o bien vendido, según el resultado del siguiente proceso: se escogen al azar dos motores y se inspeccionan. Si uno o más son defectuosos, el lote es rechazado; de otro modo es aceptado. Supóngase que cada uno de los motores cuesta $75 y se vende por $100; si el lote contiene un motor defectuoso,

¿cuál es la utilidad esperada por el fabricante? Hallar la esperanza del número de motores defectuosos encontrados. Hallar la varianza y desvío estándar de la utilidad y del número de motores defectuosos. Graficar las distribuciones y los parámetros esperanza y desvío. Hallar la asimetría del número de motores defectuosos y de la utilidad.

Este ejercicio pide el cálculo de una esperanza relativa a la ganancia bajo dos opciones distintas: perder los $75 invertidos (-75) o bien ganar $25 (+25) si se vende por $100 tras haber invertido $75. Nuestra variable aleatoria es discreta, suele decirse “dicotómica” cuando hay sólo dos alternativas posibles dada por X={-75;+25}. Falta calcular las probabilidades asociadas.
Para ello observamos que hay 10 motores y 1 es defectuoso. El poco escrupuloso vendedor lo mezcla con los 9 buenos y espera que no lo detecte en la inspección aleatoria de 2 motores. Para el vendedor será favorable que saque 2 motores buenos cualesquiera de los 9 buenos, aunque los casos posibles es que saque 2 motores cualesquiera de los 10. Luego $cp = \binom{10}{2}$ y favorables para el vendedor $cf = \binom{9}{2}$. La probabilidad de tener éxito en la venta y ganar $75 está dada por

$$P(Ganar) = P(X = 25) = \frac{\binom{9}{2}}{\binom{10}{2}} = \frac{\frac{9!}{2!\,(9-2)!}}{\frac{10!}{2!\,(10-2)!}} = \frac{9 * 8}{10 * 9} = 0{,}8$$

Por lo tanto, a partir del complemento

$$P(Perder) = P(X = -75) = 1 - 0{,}8 = 0{,}2$$

Calculamos la esperanza

$$E(X) = 0{,}8 * P(X = 25) + (-75) * P(X = -75) = 0{,}8 * 25 - 0{,}2 * 75 = 5$$

Tiene una alta probabilidad de vender pero la ganancia esperada es pequeña. En el siguiente gráfico representamos la distribución de probabilidad y la esperanza de la ganancia.
Para hallar la varianza podemos recurrir a la definición.

$$V(X) = (25 - 5)^2 * 0{,}8 + (-75 - 5)^2 * 0{,}2 = 1600$$

La varianza dio $V(X) = \$^2 1600$. El desvío estándar vale entonces $\sigma = \$40$. Si calculamos la asimetría obtenemos

$$\gamma_1 = \frac{(25-5)^3 * 0{,}8 + (-75-5)^3 * 0{,}2}{40^3} = \frac{6400 - 102400}{64000} = -1{,}5$$

El resultado es consistente con la forma de la distribución, basta comparar con los tres dibujos típicos de las distribuciones simétrica, sesgada a izquierda o negativa, y sesgada a derecha o positiva.
Calculamos la curtosis para ejercitar el procedimiento pero nos dirá poco en relación con la definición, como veremos en la unidad siguiente.

$$\gamma_2 = \frac{(25-5)^4 * 0{,}8 + (-75-5)^4 * 0{,}2}{40^4} - 3 = \frac{128000 + 8192000}{2560000} - 3 = 0{,}25$$

3. Un cable se fabrica con un diámetro variable en un rango entre 6 mm y 8 mm que responde a una función de densidad polinómica de la forma

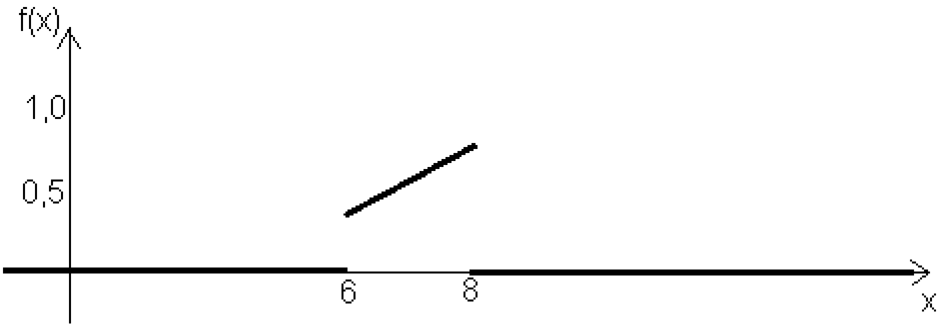
$$f_{(x)} = a(x-4)$$

a) Hallar el valor de la constante “a”.
b) Graficar la densidad y la distribución.
c) Calcular el valor esperado μ de la variable aleatoria “diámetro”.
d) Calcular la varianza y la desviación estándar σ de dicha variable aleatoria.
e) Calcular la asimetría de la variable aleatoria.
f) Calcular la probabilidad de que el diámetro sea mayor que 7.
g) Calcular la probabilidad de que el diámetro pertenezca al intervalo [6,5;7,5].
h) Indicar en los gráficos la esperanza, desvío y probabilidades calculadas.
i) Hallar la probabilidad asociada al intervalo [μ-σ; μ+σ]
j) Hallar la mediana y los cuartiles de la distribución.

Para resolver el punto a) basta recordar las dos propiedades que debe cumplir una función de densidad: no ser negativa y su integral en todo el rango debe valer uno.

$$\int_{-\infty}^{+\infty} f(x)dx = \int_6^8 a(x-4)dx = a\left(\frac{x^2}{2} - 4x\right)_6^8 = a * 6 = 1$$

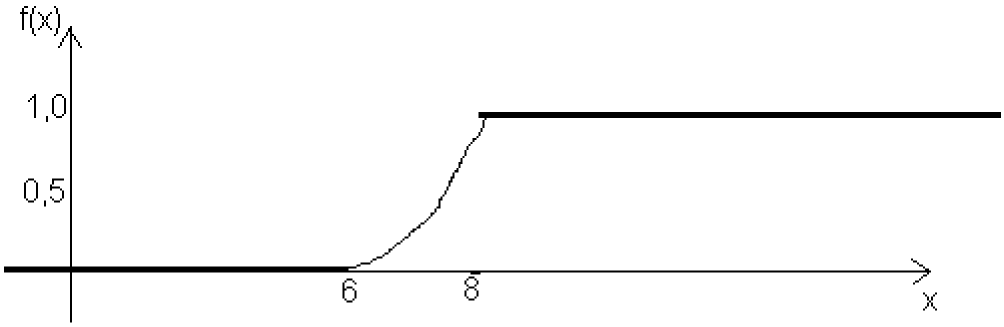
de donde a=1/6. Graficamos la función de densidad



Para poder graficar la función de distribución y completar el punto b), primero debemos hallarla, para ello calculamos la integral con el límite superior indefinido

$$F(x) = \int_{-\infty}^{x} f(x)dx = \int_6^x \frac{1}{6}(x-4)dx = \frac{1}{6}\left[\left(\frac{x^2}{2} - 4x\right) + 6\right] = \frac{x^2}{12} - \frac{2x}{3} + 1$$

Graficamos la función de distribución.



Para hallar el punto c) recurrimos a la definición de esperanza de una función continua.

$$E(x) = \mu = \int_{-\infty}^{+\infty} xf(x)dx = \int_{6}^{8} x\frac{1}{6}(x-4)dx = \frac{1}{6}\left(\frac{x^3}{3} - 4\frac{x^2}{2}\right)_6^8 = 7,\hat{1}$$

Para resolver el punto d) podemos recurrir a la definición

$$V(x) = \sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx = \int_{6}^{8} (x-\mu)^2 \frac{1}{6}(x-4)dx$$

O bien recurrir al cálculo del momento de segundo orden

$$V(x) = \sigma^2 = E(x^2) - E^2(x) = \alpha_2 - \mu^2 = \int_{6}^{8} x^2 \frac{1}{6}(x-4)dx - \mu^2$$

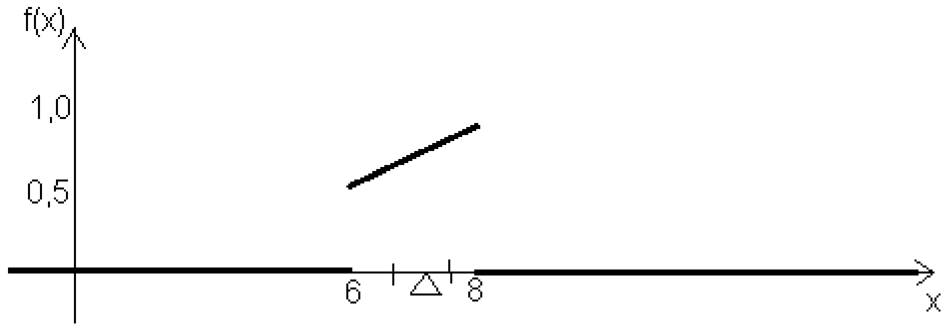
Utilizando el momento de segundo orden

$$V(x) = \sigma^2 = 50,\hat{8} - 7,\hat{1}^2 \cong 0,321$$

Con lo cual

$$\sigma = 0,567$$

Entre $\mu - \sigma \cong 6,54$ y $\mu + \sigma \cong 7,68$ tenemos casi el 60% de la distribución. En el gráfico que sigue representamos aproximadamente la posición de la esperanza y de los límites indicados por los desvíos.



Para hallar la asimetría en el punto e) podemos calcular el momento centrado de tercer orden

$$\mu_3(x) = \int_{6}^{8} (x-\mu)^3 \frac{1}{6}(x-4)dx$$

Y luego dividirlo por σ^3, o bien calcular el momento no centrado de tercer orden

$$\alpha_3 = E(x^3) = \int_{6}^{8} x^3 \frac{1}{6}(x-4)dx$$

Y recordar que

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

Con lo cual

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = -0,23$$

Para resolver el punto f) basta calcula

$$P(X > 7) = \int_7^8 \frac{1}{6}(x-4)dx = 1 - F(7) \cong 0{,}583$$

De la misma manera para el punto g)

$$P(6{,}5 < X \leq 7{,}5) = \int_{6,5}^{7,5} \frac{1}{6}(x-4)dx = F(7{,}5) - F(6{,}5) = 0{,}5$$

Para el punto h) calculamos

$$P(\mu - \sigma < X \leq \mu + \sigma) = \int_{\mu-\sigma}^{\mu+\sigma} \frac{1}{6}(x-4)dx = F(\mu+\sigma) - F(\mu-\sigma) \cong 0{,}6$$

Para hallar la mediana en i)

$$\int_6^{\tilde{x}} \frac{1}{6}(x-4)dx = 0{,}5 \qquad \tilde{x} = x_{c2} = 7{,}16$$

Y para los cuartiles

$$\int_6^{c_1} \frac{1}{6}(x-4)dx = 0{,}25 \quad x_{c1} = 6{,}65 \quad y \quad \int_6^{c_3} \frac{1}{6}(x-4)dx = 0{,}75 \quad x_{c3} = 7{,}6$$

Al disponer de los cuartiles podemos comparar la asimetría con las medidas de forma que se puede obtener comparándolos. Notemos que $x_{c2} - x_{c1} = 7{,}16 - 6{,}65 = 0{,}51$ mientras que $x_{c3} - x_{c2} = 7{,}60 - 7{,}16 = 0{,}44$. Como el intervalo entre los dos primeros cuartiles es mayor que el observado entre el segundo y el tercero, es un indicio de que la distribución tiene probabilidad más dispersa a izquierda de la mediana que a derecha, y que la asimetría es negativa. También podemos observar que la esperanza (7,11) es menor que la mediana (7,16), otro indicador de asimetría negativa.

4. Supóngase que X está distribuida uniformemente en [0,a]. Encontrar la esperanza μ y varianza σ² de X. Hallar la probabilidad de que la variable aleatoria se encuentre en el intervalo [μ-σ; μ+σ]

Vamos a resolver el problema de un modo más general, en un intervalo [a;b] cualquiera. La constante, ya lo sabemos, vale $\frac{1}{b-a}$, la esperanza se obtiene de

$$\mu = E(X) = \int_a^b x\frac{1}{b-a}dx = \frac{1}{b-a}\frac{x^2}{2}\Big|_a^b = \frac{1}{b-a}\frac{1}{2}(b^2-a^2) = \frac{1}{b-a}\frac{1}{2}(b-a)(b+a)$$
$$= \frac{a+b}{2}$$

Para calcular la varianza podemos recurrir al momento de segundo orden

$$\alpha_2 = E(X^2) = \int_a^b x^2\frac{1}{b-a}dx = \frac{1}{b-a}\frac{x^3}{3}\Big|_a^b = \frac{1}{3}\frac{1}{b-a}(b^3-a^3) = \frac{b^2+ab+a^2}{3}$$

Luego

$$\mu_2 = \sigma^2 = V(X) = E(X^2) - E^2(X) = \alpha_2 - \mu^2 = \frac{b^2+ab+a^2}{3} - \left(\frac{a+b}{2}\right)^2$$

Desarrollando

$$\sigma^2 = \frac{b^2+ab+a^2}{3} - \frac{a^2+2ab+b^2}{4} = \frac{4b^2+4ab+4a^2-3a^2-6ab-3b^2}{12}$$

Simplificando

$$\sigma^2 = \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12}$$

De donde

$$\sigma = \frac{b-a}{2\sqrt{3}}$$

En el ejercicio original $a = 0$ y $b = a$, de modo que $\mu = \frac{a}{2}$, $\sigma^2 = \frac{a^2}{12}$ y $\sigma = \frac{a}{2\sqrt{3}}$

5. Supóngase que un instrumento electrónico tiene una duración X (en unidades de 1000 horas) que se considera como una variable aleatoria continua con la siguiente función de densidad:

$$f_{(x)} = e^{-x} \;\; si\; x > 0$$

Suponiendo que el costo de fabricación de tal artículo es $2.00. El fabricante vende el artículo por $5.00, pero garantiza un reembolso total si $X \leq 0{,}9$. ¿Cuál es la utilidad esperada del fabricante por artículo? Hallar la varianza de la utilidad, el desvío estándar y la asimetría. Graficar la distribución de la duración y de la utilidad.

Este ejercicio tiene similitudes con el ejercicio 2, se pide la esperanza, varianza, desviación estándar y asimetría de una distribución discreta con dos valores posibles de ganancia (+3) y de pérdida (-2) medida en pesos por unidad. El criterio de ganancia o pérdida depende del reembolso que el cliente pueda pedir si no cumple con la garantía de duración. Esta probabilidad se obtiene de la distribución continua de duración dada por $f_{(x)} = e^{-x}$ para $x > 0$. Si $X \leq 0{,}9$ medido en horas de funcionamiento, habrá pérdida porque no cumple con la garantía y, si $X > 0{,}9$, habrá ganancia, siendo ambas probabilidades complementarias a la unidad. Calculemos

$$P(X > 0{,}9) = \int_{0,9}^{\infty} e^{-x}dx = -e^{-x}|_{0,9}^{\infty} = 0 - (-e^{-0,9}) = e^{-0,9} \cong 0{,}406$$

Luego

$$P(X \leq 0{,}9) = 1 - 0{,}406 = 0{,}594$$

Con estos valores de probabilidad podemos calcular la esperanza y la varianza de la distribución discreta asociada a la utilidad "*u*"

$$\mu = E(u) = -2 * 0{,}594 + 3 * 0{,}406 = 0{,}03$$

La varianza

$$\sigma^2 = V(u) = (-2-0{,}03)^2 * 0{,}594 + (3-0{,}03)^2 * 0{,}406 = 6{,}0291$$

Por lo tanto

$$\sigma = \sqrt{\sigma^2} = \sqrt{6{,}0291} = 2{,}455$$

Y asimetría

$$\gamma_1 = \frac{(-2-0{,}03)^3 * 0{,}594 + (3-0{,}03)^3 * 0{,}406}{2{,}455^3} = 1{,}055$$

Gráficamente

6. Cierta aleación se forma al combinar la mezcla fundida de dos metales. La aleación que resulta contiene cierto porcentaje de plomo *X*, que puede considerarse como una variable aleatoria. Supongamos que *X* tiene la siguiente función de densidad:

$$f_{(x)} = \frac{3}{5}10^{-5}x(100-x) \quad para\ 0 \leq x \leq 100$$

Suponemos que *P* es la utilidad neta obtenida al vender esta aleación (por kg) y está dada por la siguiente función del porcentaje del contenido de plomo $P = C_1 + C_2X$, en donde $C_1 = 100$ y $C_2 = 2$. Graficar la densidad, hallar y graficar la distribución del contenido de plomo, calcular su esperanza, varianza, desvío, mediana, moda y asimetría. Calcular la utilidad esperada y varianza (por kg).

El gráfico de la función de densidad asociada al contenido de plomo es un arco de parábola con raíces en $x = 0$ y $x = 100$. Es simétrica con máximo en $x = 50$. Por simetría la moda debe coincidir con la mediana y con la esperanza mientras que la asimetría tiene que ser nula.
La varianza la obtenemos del momento de segundo orden

$$\alpha_2 = \int_0^{100} x^2 \frac{3}{5}10^{-5}x(100-x)dx = \frac{3}{5}10^{-5}\left(100\int_0^{100} x^3dx - \int_0^{100} x^4dx\right) = 3000$$

$$\sigma^2 = V(X) = \alpha_2 - \mu^2 = 3000 - 50^2 = 500$$
$$\sigma = \sqrt{500} = 22{,}4$$

Para calcular la esperanza y varianza de la utilidad recurrimos a la transformación

$$E(P) = E(100 + 2X) = 100 + 2E(X) = 100 + 2 * 50 = 200$$
$$V(P) = V(100 + 2X) = 2^2V(X) = 4 * 500 * 2000$$

7. Suponiendo que X tiene una función de densidad:

$$f_{(x)} = \frac{x-4}{6} \quad para\ 6 < x < 8$$

Sea $W = \frac{1}{3}X$

a) Calcular $E_{(W)}$ usando la función de densidad de *X*.

b) Calcular $E_{(W)}$ usando la función de densidad de *W*.

Primero calculamos la esperanza de X

$$E(X) = \int_6^8 x\frac{x-4}{6}dx = \frac{1}{6}\int_6^8 (x^2 - 4x)dx = \frac{1}{6}\left(\frac{x^3}{3} - 2x^2\right)_6^8 = 7,\hat{1}$$

Para resolver el punto "a" recurrimos a la esperanza de una transformación lineal

$$E(W) = E\left(\frac{1}{3}X\right) = \frac{1}{3}E(X) = 2{,}37$$

Para resolver el punto "b" debemos hallar primer la función de densidad de W. Para ello necesitamos primero la función de distribución de W.

$$F(x) = \int_6^x \frac{x-4}{6}dx = \frac{1}{6}\left(\frac{x^2}{2} - 4x\right)_6^x = \frac{1}{12}x^2 - \frac{2}{3}x - 1$$

$$G(w) = P(W \leq w) = P\left(\frac{1}{3}X \leq w\right) = P(X \leq 3w) = F(3w) = \frac{1}{12}(3w)^2 - \frac{2}{3}(3w) - 1$$
$$= \frac{3}{4}w^2 - 2w - 1 \quad 2 \leq w \leq \frac{8}{3}$$

Si derivamos la función de distribución obtenemos la función de densidad

$$g(w) = G'(w) = \frac{3}{2}w - 2 \qquad 2 \leq w \leq \frac{8}{3}$$

Luego calculamos la esperanza

$$E(w) = \int_2^{8/3} w\left(\frac{3}{2}w - 2\right)dw = \frac{1}{2}w^3 - w^2\Big|_2^{8/3} = 2{,}37$$

8. Supóngase que X es una variable aleatoria para la cual $E_{(X)} = 10$ y $V_{(X)} = 25$. Si se define la variable aleatoria $Y = aX + b$, hallar la relación entre la esperanza y varianza de Y y de X. ¿Para qué valores positivos de a y b tiene $Y = aX + b$ esperanza 0 y varianza 1? (Se dice que la variable Y está estandarizada).

Nos apoyamos en la esperanza y la varianza aplicada a una transformación lineal con parámetros a determinar.

$$0 = E(Y) = E(aX + b) = aE(X) + b = a * 10 + b$$
$$1 = V(Y) = V(aX + b) = a^2V(X) = a^2 * 25$$

De allí

$$a = \frac{1}{5} \qquad b = -2$$

Luego

$$Y = \frac{1}{5}X - 2 = \frac{x - 10}{5}$$

A esta última expresión se la conoce como "estandarización" de una variable aleatoria, que consiste en ajustar los coeficientes lineal e independiente de modo que la esperanza de la nueva variable aleatoria sea nula y su varianza sea unitaria. La forma general de estandarizar una variable aleatoria consiste en restar la esperanza y dividir la desviación con respecto a la esperanza sobre el desvío estándar.

Unidad 5: Distribuciones de Probabilidad

Hemos visto que la importancia del conocimiento de la función de distribución $F(x)$ radica en que contiene toda la información necesaria acerca del comportamiento aleatorio del sistema bajo estudio. En los sistemas típicos, hay algunas distribuciones que se presentan de manera usual. Entre ellas hemos mencionado la distribución Binomial o de Bernoulli como una distribución discreta, y la distribución uniforme como continua. En esta unidad estudiaremos otras distribuciones de probabilidad de uso frecuente en la práctica. Antes enunciaremos un teorema muy general, de escaso interés práctico, pero en el que se fundamentan muchas demostraciones generales

Teorema de Tchebyshev

El teorema de Tchebyshev da una idea de la interpretación de la desviación como medida de dispersión en una distribución, y una primera acotación de la probabilidad de eventos extremos.
Puede expresarse diciendo que si $\mu = E(x)$ y $\sigma^2 = V(x)$ entonces

$$P(|X-\mu| > k\sigma) < \frac{1}{k^2}$$

Este teorema no exige ningún requisito especial sobre la distribución de probabilidad. Es obviamente cierto si $k < 1$. Si $k = 1$ o si $k > 1$ entonces puede plantearse que si x está fuera de los límites $(\mu - k\sigma; \mu + k\sigma)$ debe ser $(x-\mu)^2 > k^2\sigma^2$. Ahora bien, por definición es cierto que $\sigma^2 = \int_{-\infty}^{+\infty}(x-\mu)^2 f(x)dx$. Esta integral puede separarse en tres partes integrando en los intervalos $(-\infty; \mu - k\sigma)$, entre $(\mu - k\sigma; \mu + k\sigma)$ y entre $(\mu + k\sigma; +\infty)$.

$$\sigma^2 = \int_{-\infty}^{\mu-k\sigma}(x-\mu)^2 f(x)dx + \int_{\mu-k\sigma}^{\mu+k\sigma}(x-\mu)^2 f(x)dx + \int_{\mu+k\sigma}^{+\infty}(x-\mu)^2 f(x)dx$$

La suma de las tres integrales será mayor que la suma de las dos extremas, por lo que puede escribirse que

$$\sigma^2 > \int_{-\infty}^{\mu-k\sigma}(x-\mu)^2 f(x)dx + \int_{\mu+k\sigma}^{+\infty}(x-\mu)^2 f(x)dx$$

Y por la desigualdad $(x-\mu)^2 > k^2\sigma^2$

$$\sigma^2 > k^2\sigma^2\left(\int_{-\infty}^{\mu-k\sigma} f(x)dx + \int_{\mu+k\sigma}^{+\infty} f(x)dx\right) = k^2\sigma^2 P(|X-\mu| > k\sigma)$$

La última igualdad es válida por definición de probabilidad. Luego, simplificando la varianza, se obtiene el teorema tal como fue enunciado.
Si tomamos como ejemplo una distribución uniforme en el intervalo $[0; 10]$, su esperanza vale 5, su varianza vale 25/3 y su desvío $5\sqrt{3}/3$. En consecuencia, el teorema de Tchebyshev asegura que $P\left(|X-5| > 1{,}5.5.\frac{\sqrt{3}}{3}\right) < 1/1{,}5^2 = 0.45$. Pero los límites definidos por el módulo valen 0.77 y 9.33, por lo que esta probabilidad vale estrictamente 0,154, claramente menor que la acotación impuesta por el teorema.
Es claro que la acotación es muy grosera, pero no hace ninguna hipótesis sobre la distribución en particular, de allí que su importancia radica en la validez general del teorema (si bien se demostró para una distribución continua, con un poco más de notación y recurso a sumatorias, es también válido para distribuciones discretas), que nos muestra que la

probabilidad de ocurrencia de eventos es menor en tanto más alejado está el valor de la variable aleatoria del correspondiente a la esperanza.

Teorema de Bernoulli

Hemos estudiado la distribución Binomial o de Bernoulli. Ésta consiste en n repeticiones independientes de un experimento con dos alternativas y probabilidad p de ocurrencia de una de ellas llamada de "éxito".

Los parámetros de la distribución binomial son $E(x) = np$ y $V(x) = np(1-p)$. También puede probarse que la asimetría vale $A(x) = [(1-p)-p]/\sqrt{np(1-p)}$, y la curtosis resulta $C(x) = \big(1-6p(1-p)\big)/np(1-p)$.

Si en n repeticiones de un experimento binomial se obtiene X éxitos, se llama $f_r = f/n$ a la frecuencia relativa de éxitos o razón frecuencial. La esperanza será dada por el desarrollo $E(f_r) = E(X/n) = E(X)/n = np/n = p$. Por otra parte la varianza, recordando la dependencia cuadrática en la relación lineal $V(f_r) = np(1-p)/n^2 = p(1-p)/n$. Si cambiamos de denominación a $\varepsilon = k\sigma = k\sqrt{p(1-p)/n}$ resulta $k = \varepsilon\sqrt{n/p(1-p)}$.

De acuerdo con esto y el Teorema de Tchebyshev queda

$$P\left(\left|\frac{f}{n} - p\right| > \varepsilon\right) < \frac{1}{k^2} = \frac{p(1-p)}{n\varepsilon^2} < \frac{1}{4n\varepsilon^2} \xrightarrow{n\to\infty} 0$$

De allí se deduce que la frecuencia relativa tiende a la probabilidad de éxito sobre un evento cuando el número de repeticiones del experimento aleatorio tiende a infinito. El carácter asintótico de la frecuencia relativa, que tiende hacia la probabilidad, justifica que este límite pueda ser utilizado como una definición de probabilidad más general y práctica que la relativa a los juegos de azar como casos favorables y posibles, o bien como cociente de cardinales. Es una definición que se apoya en la condición de límite de sucesiones, y que habilita la aplicación práctica del cálculo de probabilidades a través de la realización concreta de pruebas repetidas. Esto nos conducirá luego a fundamentar la estadística como medio de aproximación o la estimación de valores de probabilidad válidos en la práctica.

Distribución de Poisson

Un caso límite de la distribución binomial se obtiene cuando se asume que el número de repeticiones independientes es grande (n tiende a infinito) y que la probabilidad de éxito es pequeña (p tiende a cero). Si se parte de la expresión de la distribución binomial

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

con algunas transformaciones resulta

$$P(X = k) = \frac{n!}{k!\,(n-k)!}(p)^k \frac{(1-p)^n}{(1-p)^k}$$

O bien

$$P(X = k) = \frac{n(n-1)\ldots(n-k+1)}{k!}\frac{(np)^k}{n^k}\frac{(1-p)^n}{(1-p)^k}$$

De donde

$$P(X = k) = \frac{(np)^k}{k!}\frac{(1-p)^n}{(1-p)^k}\frac{n(n-1)\ldots(n-k+1)}{n^k}$$

Luego

$$P(X = k) = \frac{(np)^k}{k!}\left(1-\frac{np}{n}\right)^n \frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{k-1}{n}\right)}{(1-p)^k}$$

Si se asume que un valor $\lambda = np$ es un número pequeño no mayor que 3 o 4 (la aproximación es válida para cualquier λ, pero sólo es útil si λ es pequeño), resulta que cuando n tiende a infinito y p tiende a cero, el último cociente tiende a 1 y

$$P(X = k) \xrightarrow[]{\substack{n\to\infty \\ p\to 0}} \frac{\lambda^k}{k!}\left(1 - \frac{\lambda}{n}\right)^n$$

Pero la expresión $\left(1 - \frac{\lambda}{n}\right)^n$ tiende a $e^{-\lambda}$ cuando n tiende a infinito. Por lo tanto puede resumirse que

$$P(X = k) \xrightarrow[]{\substack{n\to\infty \\ p\to 0}} \frac{\lambda^k}{k!} e^{-\lambda}$$

Esta expresión se conoce como “distribución de Poisson”, muy útil para el estudio de los eventos raros o poco frecuentes.

Es fácil mostrar que la suma de $P(X = k)$ sobre $k = 0$ hasta infinito vale 1 a partir del desarrollo de Taylor de e^{λ}, lo que justifica que sea una distribución de probabilidad.

Los parámetros de la distribución de Poisson son $E(X) = \lambda$ y $V(X) = \lambda$.

$$E(X) = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} = \sum_{r=0}^{\infty} \frac{\lambda^{r+1}}{r!} e^{-\lambda} = \lambda$$

En la última igualdad hemos reemplazado $r = k - 1$, y tras extraer λ fuera de la sumatoria y observar que se obtiene nuevamente la distribución de Poisson, cuya serie da por resultado la unidad, obtenemos la esperanza buscada.

Para la varianza

$$V(X) = \sum_{k=0}^{\infty} (k-\lambda)^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} - \lambda^2 = \lambda \sum_{r=0}^{\infty} (r+1) \frac{\lambda^r}{r!} e^{-\lambda} - \lambda^2 =$$

$$= \lambda \sum_{r=0}^{\infty} r \frac{\lambda^r}{r!} e^{-\lambda} + \lambda \sum_{r=0}^{\infty} \frac{\lambda^r}{r!} e^{-\lambda} - \lambda^2 = \lambda^2 \sum_{s=0}^{\infty} \frac{\lambda^s}{s!} e^{-\lambda} + \lambda - \lambda^2 = \lambda$$

La última igualdad se obtuvo de reemplazar $s = r - 1$ y observar que se obtiene nuevamente una distribución de Poisson.

Es claro que la distribución de Poisson es discreta sobre una variable aleatoria que incluye el cero y los naturales. El valor más probable o moda de la distribución se obtiene entonces comparando un k_m o valor de k con probabilidad máxima, con las probabilidades correspondientes a $k_m - 1$ y a $k_m + 1$. Es decir que sobre la condición de máximo

$$\frac{\lambda}{k_m} \frac{\lambda^{k_m - 1}}{(k_m - 1)!} e^{-\lambda} = \frac{\lambda^{k_m}}{k_m!} e^{-\lambda} = P(X = k_m) \geq P(X = k_m - 1) = \frac{\lambda^{k_m - 1}}{(k_m - 1)!} e^{-\lambda}$$

y

$$\frac{\lambda^{k_m}}{k_m!} e^{-\lambda} = P(X = k_m) \geq P(X = k_m + 1) = \frac{\lambda^{k_m + 1}}{(k_m + 1)!} e^{-\lambda} = \frac{\lambda}{k_m + 1} \frac{\lambda^{k_m}}{k_m!} e^{-\lambda}$$

Simplificando y despejando la moda, se obtiene la acotación entera

$$\lambda - 1 \leq k_m \leq \lambda$$

La moda es el entero acotado entre los dos límites propuestos, pero si el parámetro es entero, habrá una “degeneración modal” en el sentido que dos valores de *k* contiguos tienen el mismo valor máximo de probabilidad.

Distribución Normal o de Gauss

Así como hemos hallado una expresión para una distribución binomial de eventos “raros”, el teorema de De Moivre desarrolla un análisis de una condición límite sobre la distribución binomial a partir de un valor de p fijo, pero no necesariamente pequeño (en condiciones de máxima incertidumbre $p = 1/2$) y el número de repeticiones que tiende a infinito.

A partir de la expresión de la distribución binomial, y definiendo una variable estandarizada $\lambda = (f - np)/\sqrt{np(1-p)}$, resulta que la esperanza $E(\lambda) = 0$ y la varianza $V(\lambda) = 1$. Por lo tanto la probabilidad de que esta variable esté entre dos límites vale $P(\lambda_1 \leq \lambda \leq \lambda_2)$ y es equivalente a la probabilidad de que el número de éxitos en un experimento binomial con *f* éxitos se encuentre entre los límites

$$f_1 = np + \lambda_1\sqrt{np(1-p)}$$

y

$$f_2 = np + \lambda_2\sqrt{np(1-p)}$$

lo cual es igual a la suma de los términos binomiales para *f* entre estos límites.

$$P(\lambda_1 \leq \lambda \leq \lambda_2) = P(f_1 \leq f \leq f_2) = \sum_{f=f_1}^{f=f_2} \binom{n}{f} p^f (1-p)^{n-f}$$

Si luego tomamos un número de repeticiones *n* que crece indefinidamente mientras que la probabilidad de éxito *p* se mantiene constante, el número de términos en la suma tiende a infinito pero no la suma de los términos.

Haciendo una serie de aproximaciones y acotaciones queda

$$P(\lambda_1 < \lambda \leq \lambda_2) = \frac{1}{\sqrt{2\pi np(1-p)}} \sum_{\lambda=\lambda_1}^{\lambda=\lambda_2} e^{-\lambda^2/2}$$

En el límite para n tendiendo a infinito, esta expresión tiende a una integral de la forma

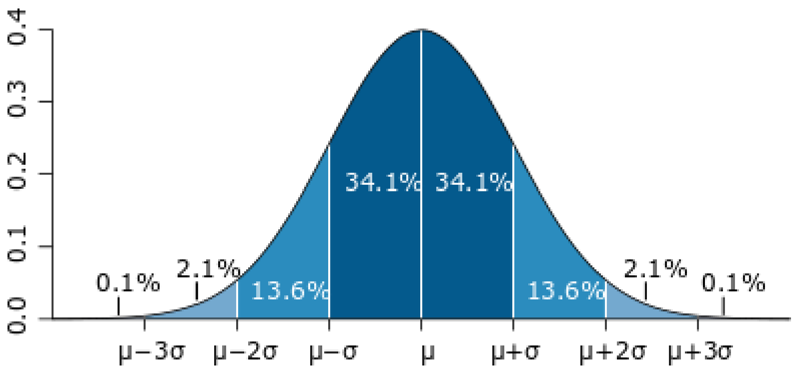
$$P(\lambda_1 < \lambda \leq \lambda_2) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\lambda_1}^{\lambda_2} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

donde $\sigma^2 = np(1-p)$ y $\mu = np$.

Más allá de los detalles omitidos de una demostración algebraicamente extensa, puede verse que en el límite para *n* tendiendo a infinito, la distribución binomial, discreta, tiende a la forma de una distribución continua conocida como “distribución normal o de Gauss”, también mencionada como “campana de Gauss” por la forma del gráfico de su función de densidad.

A una distribución normal se la simboliza $N(\mu; \sigma^2)$ de una manera general. Por ejemplo, $N(3; 4)$ refiere una distribución gaussiana con esperanza 3 y varianza 4, por lo que la desviación estándar vale 2.

Uno de los aspectos más relevantes que le dan la importancia que tiene, es que describe el comportamiento de una variable aleatoria continua de un experimento que se realiza totalmente al azar. Sólo hay dos parámetros que determinan la forma de la función de densidad y del cálculo de probabilidad a través de la distribución: la esperanza (μ) y la varianza (σ^2). En la siguiente figura vemos un gráfico típico

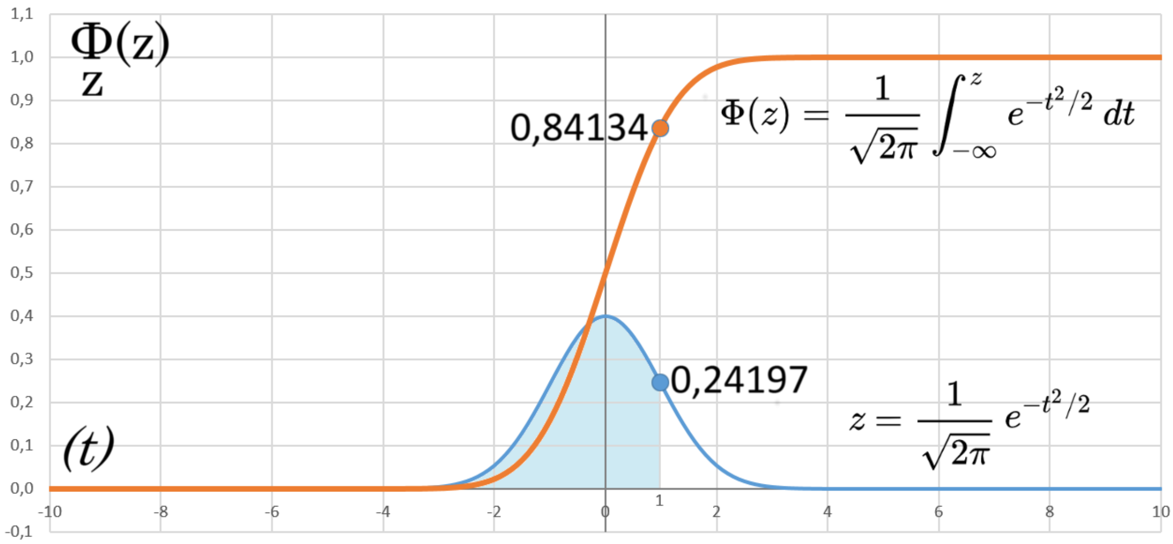


https://es.wikipedia.org/wiki/Distribución_normal

En el gráfico vemos que los puntos de inflexión coinciden con la esperanza sumado y restado un desvío estándar ($\mu \pm \sigma$), mientras que a una distancia de tres desviaciones la función de densidad prácticamente es nula, si bien formalmente tiende a cero en ambos extremos del eje real.

Para representar gráficamente una distribución normal basta observar que tiene una sola moda y es simétrica, de modo que la esperanza, mediana y moda coinciden en el punto medio de la gráfica. Los puntos de inflexión de la "campana" coinciden con los valores numéricos de $\mu \pm \sigma$. La probabilidad correspondiente a la región ubicada entre la esperanza y un desvío estándar a izquierda y derecha (entre μ-σ y μ+σ) es de 68,2%. En los límites $\mu \pm 2\sigma$ está contenido más del 95% de la distribución y dentro del intervalo $\mu \pm 3\sigma$, más del 99% de la distribución. De allí que tener eventos más alejados que tres desvíos con respecto a la normal son muy raros.

Presentamos ahora conjuntamente la función de densidad y la función de distribución.



https://es.wikipedia.org/wiki/Tabla_normal_estándar

En azul podemos ver el gráfico de la función de densidad normal y el área sombreada en celeste, que representa la probabilidad acumulada. En naranja podemos ver la función de distribución, que va progresivamente calculando el área encerrada bajo la curva de la función de densidad interpretándola como la probabilidad acumulada correspondiente a cada punto. En el gráfico se ha indicado un valor como ejemplo de 0,84134, correspondiente a $t = 1$ con valor de densidad $\varphi(t) = 0{,}24197$.

Desde el punto de vista del cálculo, el problema que presenta la distribución normal es que su función de densidad no es integrable analíticamente, es decir, no tiene una primitiva, y no puede aplicarse la regla de Barrow. Por tal motivo, no es posible hallar valores exactos de probabilidad y en la práctica se recurre a estimaciones numéricas que se presentan en tablas de la distribución normal. En estas tablas se dispone de valores numéricos de una única distribución normal conocida como "estándar" con esperanza cero y varianza uno, es decir $N(0;1)$. Es usual que se indique $\varphi(z)$ como una referencia a la función de densidad normal estándar, y $\Phi(z)$ como una forma sintética de expresar la función de distribución normal estándar.

Para utilizar estas tablas en distribuciones normales con otros valores de μ y σ, se estandariza la variable aleatoria original X por medio de la relación de cambio de variable

$$Z = \frac{X-\mu}{\sigma}$$

Considerando

$$P(x_1 < x \le x_2) = \frac{1}{\sqrt{2\pi}\sigma}\int_{x_1}^{x_2} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Utilizando el mencionado cambio de variable,

$$z_1 = \frac{x_1-\mu}{\sigma} \qquad z_2 = \frac{x_2-\mu}{\sigma}$$

de lo que resulta

$$dz = \frac{dx}{\sigma}$$

Y queda

$$P(x_1 < x \le x_2) = \frac{1}{\sqrt{2\pi}\sigma}\int_{x_1}^{x_2} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{\sqrt{2\pi}}\int_{z_1}^{z_2} e^{-\frac{1}{2}z^2} dz = P(z_1 < x \le z_2)$$

Presentamos a continuación un formato de la tabla de distribución normal.

Tabla de valores de probabilidad acumulada (ϕ) para la Distribución Normal Estándar

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3	0.0013	0.0010	0.0007	0.0005	0.0003	0.0002	0.0002	0.0001	0.0001	0.0000
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3	0.9987	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

1. Si una variable normal X no es estándar, entonces sus valores deben ser estandarizados mediante la transformación: Z=(X-μ)/σ es decir, P(X<x)=ϕ[x(-μ)/σ]
2. Para valores de z>4, ϕ[z]=1, a una precisión de cuatro decimales; para valores de z<-4, ϕ[z]=0, con cuatro decimales sigificativos.
3. Aquellos valores al lado del valor de 3 corresponden a las probabilidades acumuladas de z igual a 3.0, 3.1, 3.2, etc.

https://es.calameo.com/read/00500660270f7c1129194

Estas tablas contienen entonces la distribución de la variable aleatoria estandarizada $\Phi(z) = P(Z \leq z)$. Para utilizar la tabla de distribución normal acumulada estándar tomamos como referencia la esperanza y la varianza indicadas, y el o los límites dentro de los cuales se quiere calcular la probabilidad pedida.

Si buscamos $P(X \leq a) = P(Z \leq z_a) = \phi(z_a)$, para ello estandarizamos el valor "a" en la forma

$$z_a = \frac{a-\mu}{\sigma}$$

Con este valor numérico vamos a la tabla. Buscamos el entero y el primer decimal en la columna derecha, si es negativo, en el lado izquierdo de la tabla y, si es positivo, en el lado derecho. El segundo decimal lo buscamos como columna en la que corresponda para obtener la probabilidad buscada en la intersección de la fila y la columna. Si, por ejemplo, el valor estandarizado fuese $z_a = -1{,}26$, buscamos en la fila "-1,2" hasta la columna "0,06" para hallar 0,1038 como probabilidad deseada.

Como ejemplo más completo, si partimos de $X \sim N(3{,}4)$ se obtiene $Z = (X-3)/2$ con $Z \sim N(0{,}1)$ donde $\Phi(z) = F(z)$ está disponible en tablas. Así, si buscamos la probabilidad $P(X \leq 1.5)$ será $P(X \leq 1.5) = P(Z \leq (1{,}5-3)/2) = P(Z \leq -0{,}75)$, valor que se obtiene en tablas. Buscamos el valor de probabilidad en la fila precedida por "$-0{,}7$" y en la columna encabezada por "0,05", obteniendo el número "0,2266". Por otra parte, si se busca el valor de "z" a partir de la probabilidad haciendo uso de la tabla en sentido inverso, se obtiene el valor de "x" en la forma $x = \mu + z\sigma$.. Por ejemplo, si se quiere el valor de "z" que delimita el 97,5% de probabilidad, buscamos en la tabla el valor 0,9750 y hallamos que corresponde a $z = 1{,}96$. Es posible que no encontremos un número exacto. Por ejemplo, si buscamos el valor de "z" que acumula el 95% de probabilidad, veremos que con $z = 1{,}64$ no alcanza y que con $z = 1{,}65$ excede la probabilidad deseada en la misma proporción. Interpolando linealmente entre ambos límites y valores, obtenemos $z = 1{,}645$ como valor más aproximado que acumula el 95% de probabilidad.

En caso que se pida que sea mayor que cierto valor, recurrimos al complemento, así

$$P(X > a) = 1 - P(X \leq a) = 1 - P(Z \leq z_a) = 1 - \phi(z_a) = 1 - 0{,}1038 = 0{,}8962$$

Y si se busca una probabilidad entre dos límites

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a)$$

estandarizamos ambos límites de modo que

$$z_a = \frac{a-\mu}{\sigma} \qquad y \qquad z_b = \frac{b-\mu}{\sigma}$$

Luego

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = P(Z \leq z_b) - P(Z \leq z_a) = \phi(z_b) - \phi(z_a)$$

Mencionemos que hay diferentes presentaciones de la tabla de distribución normal. La ofrecida es estrictamente la función de distribución. En ocasiones se ofrece la cola derecha por un criterio de positividad, o sólo la mitad de la tabla dado que la otra mitad se puede construir por un criterio de simetría.

Es útil conocer los primeros momentos o parámetros de la distribución normal estándar. En el caso de la esperanza, para la distribución normal estándar vale 0 y, para una distribución normal genérica, vale μ. La varianza de la distribución normal estándar vale 1 mientras que para una normal genérica vale σ^2. La asimetría vale cero para toda distribución normal. Lo singular es que el momento centrado en la esperanza de cuarto orden dela distribución normal estándar vale 3, es decir, $E[(X-\mu)^4] = 3\sigma^4$. Por tal motivo se define el parámetro, que hemos llamado "curtosis" como

$$\gamma_2 = \frac{E[(X-\mu)^4]}{\sigma^4} - 3$$

de manera tal que al dividir por la cuarta potencia de sigma, el cuarto momento se estandarice y al restar 3 se obtenga una medida comparativa absoluta de una distribución asociada a X con respecto a la normal. Si $\gamma_2 < 0$ se dice que la distribución es “platicúrtica” o más plana que la correspondiente a la normal, mientras que si $\gamma_2 > 0$ se la llama “leptocúrtica” o más delgada y concentrada que la normal.

Hemos dicho que en todo el dominio de los reales, es decir, para una variable aleatoria definida sobre todo el eje real, la distribución normal es la de máxima aleatoriedad. Es frecuente que se utilice como hipótesis para desarrollos teóricos en los que se asume que hay aleatoriedad absoluta, también en el marco de la teoría de errores aleatorios en torno a un valor medido, en técnicas de control de calidad en tanto se supone aleatoriedad pura y por lo tanto normalidad en la distribución de alteraciones, y se impone límites de aceptación a los extremos permitidos.

Proceso Poisson

Dentro de los problemas probabilísticos se encuentra la evaluación de la probabilidad de que un suceso ocurra cierto número de veces dentro de determinado intervalo de tiempo, como el número de llamadas telefónicas en un día, el número de accidentes en un año, etc. Es decir que se busca un valor para $P(X = k/T = t)$, donde T es la variable aleatoria continua “tiempo” de duración de un proceso y X es la variable aleatoria “número de eventos” durante el tiempo indicado. Esto se interpreta como la evaluación de la probabilidad condicional de que ocurran k eventos sabiendo que el proceso ha durado cierto tiempo t. Digamos que puede generalizarse a que el condicionante sea una variable espacial, o un límite de temperaturas o lo que pudiera ser condicionante continuo de la observación de algún evento discreto. Remitiremos el desarrollo al tiempo por ser lo más usual y facilitar la comprensión evitando un grado de abstracción innecesario, pero basta ajustar la interpretación y la nomenclatura para obtener un desarrollo equivalente.

Se hará dos suposiciones idealizando el proceso. En primer lugar, que el número de eventos dentro de cierto intervalo de tiempo es independiente del número de eventos ocurridos previamente a ese intervalo. En segundo lugar, y haciendo uso de la noción de límite, se supondrá que la probabilidad de ocurrencia de un único evento en un intervalo *dt*, es decir, entre *t* y *t*+*dt*, se incremente linealmente con *dt* en primera aproximación. Por lo tanto, si λ es un parámetro fijo,

$$P(X = 1/T = dt) = \lambda dt$$

expresa la probabilidad de observar un evento $(X = 1)$ en un intervalo de tiempo diferencial $(T = dt)$ está dado por un diferencial de probabilidad (λdt). La probabilidad de que se registre más de un evento en tal intervalo *dt* es un diferencial de segundo orden y no se considera.

Para ello evaluamos primero la probabilidad de que no se presente ningún evento hasta el tiempo *t*, es decir, $P(X = 0/T = t) = p_0(t)$ presentándose el primer evento inmediatamente después de *t*, por lo tanto en *t*+*dt*. Podemos escribir esta expresión de igualdad entre intervalos reales $[0, t + dt] = [0, t] \cup (t, t + dt]$. La probabilidad de que se presente un evento en el intervalo $(t, t + dt]$ vale λdt, por lo tanto la probabilidad de que no se presente será *1*-λdt. La probabilidad de que no se presente en $[0, t + dt]$ será $p_0(t)(1 - \lambda dt)$, es decir que el evento no ocurre hasta el tiempo *t* y tampoco ocurre independientemente en el intervalo $(t, t + dt]$. En consecuencia

$$P(X = 0/T = t + dt) = p_0(t)(1 - \lambda dt) = p_0(t) - \lambda p_0(t)dt =$$
$$P(X = 0/T = t) - \lambda p_0(t)dt$$

Expresando luego que

$$P(X=0/T=d+dt)-P(X=0/T=t)=p_0(t+dt)-p_0(t)$$

o

$$d\big(p_0(t)\big)=-\lambda p_0(t)dt$$

Integrando queda

$$p_0(t)=e^{-\lambda t}+C$$

Como $p_0(0)=1$ porque es evidente que no se ha registrado ningún evento en *t=0*, resulta que *C = 0*. Luego

$$P(X=0)=p_0(t)=e^{-\lambda t}=e^{-t/\tau}$$

Esta expresión representa la distribución de tiempo durante el cual no ocurre ningún evento Esta probabilidad tiende a cero cuando el exponente tiende a infinito. Esto quiere decir que, a medida que transcurre el tiempo, la probabilidad de que tal evento no ocurra es cada vez menor. En términos complementarios, la probabilidad de que el dado evento ocurra tiende a uno. Así como

$$P(X=0)=p_0(t)=e^{-\lambda t}=e^{-t/\tau}$$

resulta

$$P(X=1)=p_1(t)=1-e^{-\lambda t}=1-e^{-t/\tau}$$

Este desarrollo puede extenderse a un número de eventos *k* no nulo en el intervalo *t+dt*

$$P(X=k/T=t+dt)=P(X=k/T=t)(1-\lambda dt)+P(X=k-1/T=t)\lambda dt$$

Esto se interpreta como expresión de *k* eventos en el tiempo *t* y ninguno en el tiempo *dt*, o bien *k-1* eventos hasta *t* y un evento más en *dt*. Podemos expresarlo

$$\frac{\big(P(X=k/T=t+dt)-P(X=k/T=t)\big)}{dt}=\frac{dP(X=k/T=t)}{dx}=$$

$$=-\lambda\,P(X=k/T=t)+\lambda P(X=k-1/T=t)$$

Esta es una ecuación diferencial con solución

$$P(X=k/T=t)=\frac{(\lambda t)^k}{k\,!}e^{-\lambda t}$$

Esta expresión se interpreta como que el número de eventos que ocurren hasta un cierto tiempo *t* es una distribución de Poisson con parámetro λt.

Como ejemplo de aplicación, puede plantearse que si se sabe que en una cierta localidad se producen dos granizadas al año, puede tratar de hallarse la probabilidad de que se produzcan exactamente 10 granizadas en cinco años o la probabilidad de que en un mes no se produzcan granizadas. Para ello se puede considerar que la probabilidad diaria de granizada vale 2/365 como valor de λ. En tal caso la probabilidad de que no se produzca una granizada en un mes se obtiene a través de la distribución exponencial como tiempo de espera hasta la primera granizada, por lo tanto $P(X=0/T=t)=e^{-\lambda t}=\ e^{-\frac{2}{365}*30}=0.85$. La probabilidad de que se produzcan 10 eventos en cinco años vale por lo tanto

$$P(X=10/=5*365)=\frac{\left(\frac{2}{365}*5*365\right)^{10}}{10!}e^{-\frac{2}{365}*5*365}=\frac{10^{10}}{10!}e^{-10}=12.5\%$$

Distribución exponencial

Destacaremos que una interpretación particular del proceso Poisson es lo que hemos planteado como "distribución exponencial" o distribución de tiempos de espera hasta la ocurrencia del primer evento. Es muy simple de plantear si se parte del proceso Poisson y se considera la probabilidad de que transcurra cierto tiempo sin que se observe la ocurrencia de ningún evento.

Si partimos de la distribución de Poisson para "k=0 dado que ha transcurrido un cierto tiempo t", resulta

$$P(X = 0/T = t) = \frac{(\lambda t)^k}{k\,!} e^{-\lambda t} = \frac{(\lambda t)^0}{0\,!} e^{-\lambda t} = e^{-\lambda t}$$

Nuestra nueva variable aleatoria propuesta es T como "tiempo de espera hasta la ocurrencia del primer evento". Como ésta es la probabilidad de que no ocurra ningún evento y nuestro proceso termina cuando $k = 1$, podemos replantear el problema escribiendo

$$P(T > t/k = 0) = e^{-\lambda t}$$

En consecuencia

$$F(t) = P(T \leq T/k = 0) = 1 - e^{-\lambda t}$$

La función de densidad se obtiene como la derivada

$$f(t) = \frac{dF(t)}{dt} = \lambda e^{-\lambda t}$$

En otros términos, podemos decir entonces que la distribución de tiempos de no ocurrencia del evento es $p_0(t)$ mientras que la distribución de probabilidad de tiempos de espera hasta la ocurrencia del primer evento es $F(t) = p_1(t)$. Llamaremos a $p_1(t)$ la "distribución exponencial", que tiende monótonamente a uno cuando el tiempo tiende a infinito y es nula para tiempos negativos ($F(t) = E_\lambda(t)$).

$$E_\lambda(t) = \begin{cases} 0 & t < 0 \\ 1 - e^{-t/\tau} & t \geq 0 \end{cases}$$

Si la derivamos, obtenemos la función de densidad asociada

$$F'(t) = \frac{dp_1(t)}{dt} = \lambda e^{-\lambda t} = \frac{1}{\tau} e^{-t/\tau}$$

Para hallar la esperanza y la varianza de la distribución exponencial se requiere una integración por partes de

$$E(T) = \int_0^\infty t\,\lambda e^{-\lambda t} = \frac{1}{\lambda}$$

y la varianza

$$V(T) = \int_0^\infty \left(t - \frac{1}{\lambda}\right)^2 \lambda e^{-\lambda t} = \frac{1}{\lambda^2}$$

de donde la desviación estándar vale

$$\sigma = \frac{1}{\lambda}$$

En general es más sencillo interpretar el tiempo esperado como la recíproca del parámetro de la distribución exponencial, dado que λ tiene unidades de recíproca de tiempo. La recíproca de λ se interpreta como "el tiempo esperado hasta la ocurrencia del primer evento", donde "tiempo esperado” refiere a la esperanza probabilística y no a una “espera” en el sentido coloquial del término.

Es frecuente que se note

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

al valor esperado del tiempo de espera hasta la ocurrencia del primer evento. La probabilidad de exceder este tiempo de espera vale

$$P(T > \mu) = P(T > \mu/k = 0) = e^{-\lambda\mu} = e^{-1} = 0{,}368$$

Una de las aplicaciones es la relativa al “tiempo de vida media”, que suele tener dos interpretaciones. Lo que hemos llamado “tiempo de espera hasta la ocurrencia del primer evento” o “tiempo esperado” o “tiempo característico”. Sin embargo esta expresión es más acorde con la idea de “mediana” en el sentido del tiempo que debe transcurrir hasta que la probabilidad de ocurrencia del primer evento vale 0,5=50%. En tal caso

$$P\left(T \le t_{\frac{1}{2}}\right) = 1 - P\left(T > t_{\frac{1}{2}}\right) = 0{,}5 = e^{-\lambda t_{1/2}}$$

de donde

$$t_{1/2} = -\frac{\ln(0{,}5)}{\lambda} = \frac{\ln{(2)}}{\lambda}$$

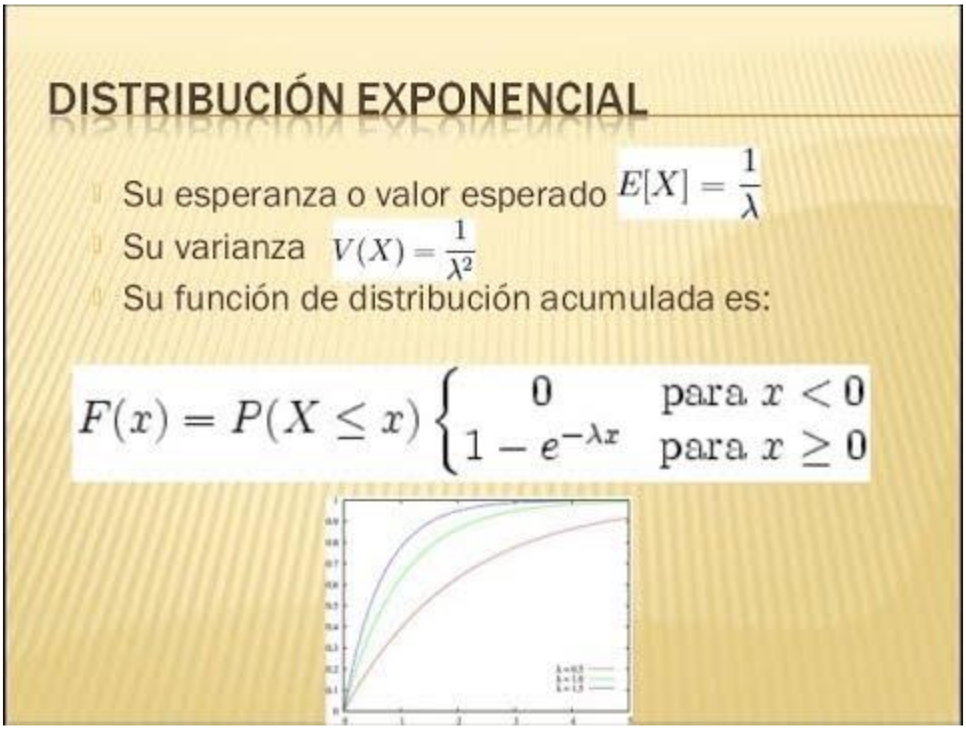
Esto es consistente con la aplicación que en física nuclear se da al tiempo que debe transcurrir para que la proporción de átomos en una muestra radioactiva se reduzca a la mitad, lo que se plantea a partir de $N = N_0 e^{-\lambda t}$ considerando que cada decaimiento es independiente de la historia previa del átomo.

Otra de las propiedades singulares es la que se llama de “falta de memoria”, que es consecuencia de las hipótesis planteadas en el proceso Poisson. Si se sabe que un evento no ha ocurrido hasta cierto tiempo t_1 y se plantea la probabilidad de que no ocurra hasta un cierto tiempo t adicional, formalmente lo escribimos

$$P(T > t_1 + t/T > t_1) = \frac{e^{-\lambda(t_1+t)}}{e^{-\lambda t_1}} = e^{-\lambda t} = P(T > t)$$

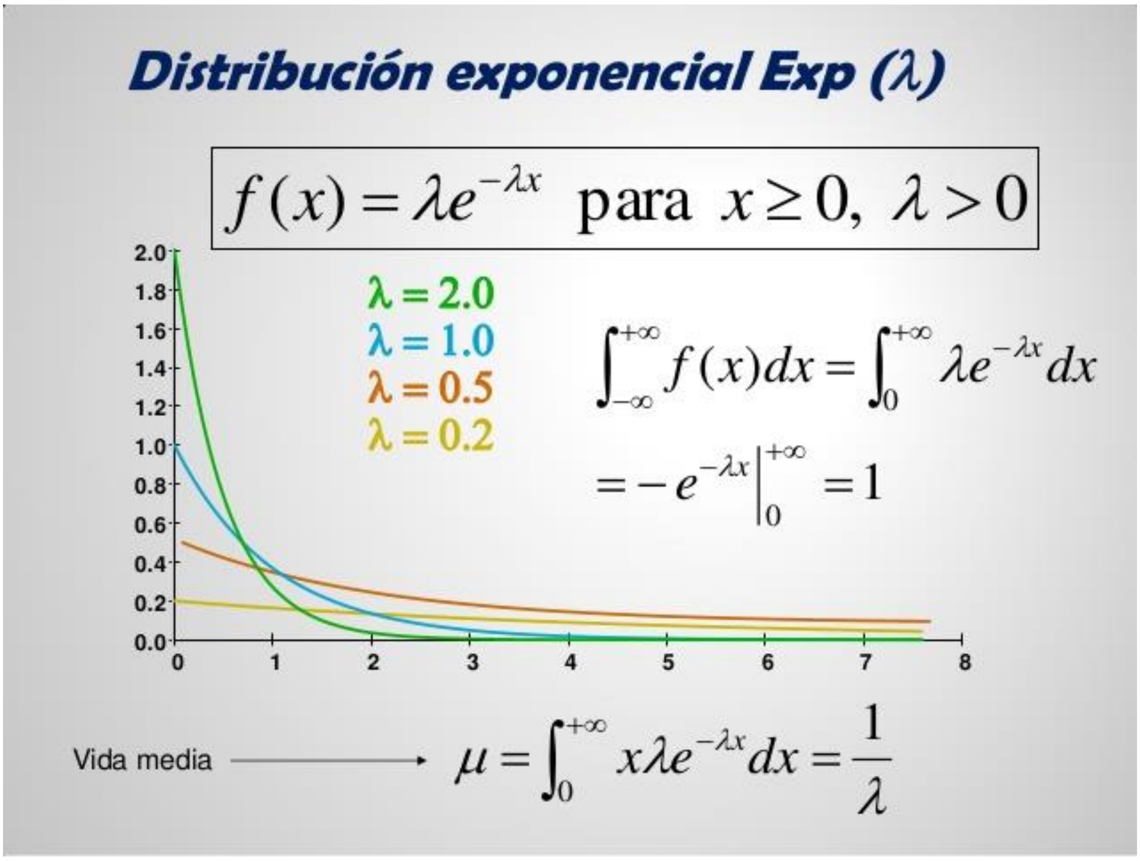
En la última expresión obtenemos que la probabilidad de que el evento no haya ocurrido hasta cierto tiempo *t* es independiente de que no haya ocurrido previamente hasta $t_1$dado que no aparece explícitamente en la expresión de la probabilidad.

El gráfico de la función de distribución es de la forma



https://www.youtube.com/watch?v=AUap_SdE5iQ

y de la de densidad



https://sites.google.com/site/probabilidad2017a/agosto

Unidad Nº 5: Funciones de Distribución

1. Un fenómeno aleatorio se produce de acuerdo con una distribución de Poisson de parámetro λ=2. Hallar la probabilidad de por lo menos dos ocurrencias del evento, de exactamente dos ocurrencias, y la esperanza y varianza de la distribución.

2. Un fenómeno especial se estima que se produce con una probabilidad diaria del 1%. Calcular la probabilidad de que se produzca por lo menos una vez en el año y de que se produzca exactamente tres veces en el año. Calcular la esperanza y varianza del número de ocurrencias.

3. Si X tiene una distribución de Poisson con parámetro λ y si $P(X = 0) = 0{,}2$, calcular $P(X > 2)$.

4. Sea X una distribución de Poisson con parámetro λ. Encontrar el valor de k para el cual $P(X = k)$ es máxima. [Indicación: comparar $P(X = k)$ con $P(X = k - 1)$].

5. Supóngase que la probabilidad de que un artículo producido por una máquina especial sea defectuoso vale 0,2. Si 10 artículos producidos se seleccionan al azar, ¿cuánto vale la probabilidad de que no se encuentre más de un artículo defectuoso? Use las distribuciones binomial y de Poisson, y compare las respuestas.

6. Una compañía de seguros ha descubierto que sólo alrededor de 0,1% de la población tiene cierto tipo de accidente cada año. Si los 10.000 seleccionados fueran elegidos en forma aleatoria entre la población, ¿cuánto vale la probabilidad de que exactamente tres de estos clientes tengan un accidente de ese tipo el próximo año? Use las distribuciones binomial y de Poisson y compare las respuestas.

7. El número de buques tanques, digamos N, que llegan cada día a cierta refinería tiene una distribución de Poisson con parámetro λ=2. Las actuales instalaciones portuarias pueden despachar tres buques al día. Si más de tres buques tanques llegan en un día, los que están en exceso deben enviarse a otro puerto.
   a) En un día determinado, ¿cuánto vale la probabilidad de tener que hacer salir buques tanques?
   b) ¿En cuánto deben aumentarse las instalaciones actuales para reducir esta probabilidad a menos del 10%?
   c) ¿Cuánto vale el número esperado de buques tanques que llegan en un día?
   d) ¿Cuánto vale el número más probable de buques tanques que llegan diariamente?
   e) ¿Cuánto vale el número esperado de buques tanques atendidos diariamente?
   f) ¿Cuánto vale el número esperado de buques tanques devueltos diariamente?

8. Supóngase que X tiene una distribución $N(2; 0{,}16)$. Use la tabla de distribución normal para evaluar las probabilidades siguientes probabilidades:
   a) $P(X \leq 2{,}1)$ b) $P(X > 1{,}8)$ c) $P\ (1{,}8 < X \leq 2{,}1)$.

9. Se sabe que un material se fabrica con un parámetro de calidad que responde a una distribución normal con esperanza μ = 20 y varianza σ² = 4.
   a) Calcular la probabilidad de que el número de ocurrencias sea menor que 17.
   b) Calcular la probabilidad de que el número de ocurrencias sea mayor que 21.
   c) Calcular la probabilidad de que el número de ocurrencias esté entre 17 y 21.
   d) Si sólo se puede desechar el 20% del material que menos responde al criterio de calidad empleado, ¿qué límites impondría a la variable aleatoria de manera tal que el parámetro no sea menor que tal límite? ¿Cuál sería el límite si se impone que el parámetro no sea mayor que cierto límite con esa probabilidad? ¿Cuáles serían los límites si se impone que el parámetro se encuentre fuera de ellos con 10% de probabilidad a cada lado?

10. El diámetro de un cable eléctrico está distribuido normalmente con esperanza 0,8 y varianza 0,0004. ¿Cuánto vale la probabilidad de que el diámetro sobrepase 0,81cm? Suponiendo que el cable se considere defectuoso si el diámetro se diferencia de su esperanza en más de 0,025, ¿cuánto vale la probabilidad de obtener un cable defectuoso?

11. Se sabe que los errores en cierto instrumento para medir longitudes están distribuidos normalmente con valor esperado cero y desviación estándar 1cm. ¿Cuánto vale la probabilidad de que al medir, los errores sean mayores que 1cm, 2cm y 3cm?

12. Suponiendo que la duración de los instrumentos electrónicos D1 y D2 tienen distribuciones N(40;36) y N(45;9) respectivamente, ¿cuál debe preferirse para usarlo durante un período de 45 horas? ¿Cuál debe preferirse para usarlo durante un período de 48 horas?

13. Un instrumento tiene una duración que responde a una distribución exponencial de parámetro 0,05 medido en meses. Calcular la probabilidad de que el instrumento dure un año, dos años y tres años. Indicar su esperanza y varianza.

14. Sea T la duración de un instrumento electrónico en meses con densidad:

$$f(t) = \lambda e^{-\lambda t} \; para \; t > 0$$

Si λ=0,02, ¿por cuántos meses de duración debe garantizar el fabricante sus instrumentos si busca que la probabilidad de que satisfaga la garantía valga 0,8?

15. Se asume que la duración de una conversación telefónica es una variable aleatoria X con función de densidad $f(x) = ae^{-\alpha x}$. Mostrar que la probabilidad de que una conversación dure más que $t_1 + t_2$ minutos, dado que ha durado por lo menos t_1 minutos, es igual a la probabilidad no condicional de que ésta dure por lo menos t_2 minutos. (Sugerencia: elaborar un ejemplo numérico y luego plantear el problema en general en términos de probabilidad condicional).

16. Compare la *cota superior* de la probabilidad $P[|X - E(X)| > 2\sqrt{V(X)}]$ obtenida con la desigualdad de Tchebychev con la probabilidad exacta en cada uno de los casos siguientes:
a) $P\big(|X - E(X)| > 2\sigma(X)\big)$ si X tiene distribución N(μ;σ²) y
b) $P\left(|X - E(X)| > \frac{2}{\lambda}\right)$ si X tiene distribución exponencial con parámetro λ.

17. La vida media de un átomo radioactivo se define como el tiempo que tarda en desintegrarse la mitad del material radioactivo. El proceso de desintegración responde a una distribución exponencial de parámetro λ ($N = N_0 e^{-\lambda t}$). Si se mide el tiempo de vida media y resulta ser de 10 años, hallar el valor de λ. ¿A qué parámetro estadístico asociaría el tiempo de vida medio? Comparar el tiempo de vida medio con la esperanza de la exponencial, interpretar esta diferencia en relación con la forma de la distribución. Hallar la probabilidad de que en un año, un mes, un día y una hora se registre una desintegración en un átomo determinado. Si la muestra contiene 10000 átomos, hallar la probabilidad de que en una hora se registren dos desintegraciones.

Unidad Nº 5: Ejercicios resueltos

1. Un fenómeno aleatorio se produce de acuerdo con una distribución de Poisson de parámetro λ=2. Hallar la probabilidad de por lo menos dos ocurrencias del evento, de exactamente dos ocurrencias, y la esperanza y varianza de la distribución.

Con el objetivo de practicar el cálculo con la distribución de Poisson, se informa los datos requeridos a través del único parámetro. Calcular la probabilidad de por lo menos dos ocurrencias requiere la suma desde k=2 hasta infinito, pero puede recurrirse al complemento en relación con k=0 y k=1.

$$P(X=0)+P(X=1)=\frac{2^0}{0!}e^{-2}+\frac{2^1}{1!}e^{-2}=3e^{-2}=0{,}406$$

Luego

$$P(X\geq 2)=1-P(X<2)=1-0{,}406=0{,}594$$

La probabilidad de exactamente dos ocurrencias se obtiene de un reemplazo directo

$$P(X=2)=\frac{2^2}{2!}e^{-2}=2e^{-2}=0{,}271$$

Recurriendo al desarrollo de la teoría, la esperanza y varianza coinciden con el valor del parámetro de la distribución de Poisson

$$E(X)=V(X)=\lambda=2$$

2. Un fenómeno especial se estima que se produce con una probabilidad diaria del 1%. Calcular la probabilidad de que se produzca por lo menos una vez en el año y de que se produzca exactamente tres veces en el año. Calcular la esperanza y varianza del número de ocurrencias.

La distribución de Poisson está asociada a eventos “raros”, poco frecuentes, lo que se sugiere al decir que se trata de un fenómeno “especial”. Si la probabilidad de ocurrencia del evento es de 1% en cada día, en términos anuales el número de días en que el experimento se repite vale 365, de modo que p=0,01 y n=365 son los parámetros de la distribución binomial que correspondería a este experimento. Pero tanto es la probabilidad pequeña como el número de repeticiones grande, de modo que estaríamos habilitados a aplicar la distribución de Poisson como condición límite y aproximación válida.

En primer lugar calculamos el parámetro que determina la distribución de Poisson

$$\lambda=np=365*0{,}01=3{,}65$$

Luego nuestra distribución es

$$P(X=k)=\frac{3{,}65^k}{k!}e^{-3{,}65}$$

Para calcular la probabilidad de que ocurra al menos una vez al año recurrimos al complemento

$$P(X\geq 1)=1-P(X=0)=1-\frac{3{,}65^0}{0!}e^{-3{,}65}=1-e^{-3{,}65}=0{,}974$$

La probabilidad de que ocurra exactamente tres veces en el año

$$P(X=3)=\frac{3{,}65^3}{3!}e^{-3{,}65}=0{,}211$$

Recordamos que la esperanza y varianza del número de ocurrencias coincide con el parámetro de Poisson, de modo que

$$E(X) = V(X) = \lambda = 3{,}65$$

3. Si X tiene una distribución de Poisson con parámetro λ y si $P(X = 0) = 0{,}2$, calcular $P(X > 2)$.

Este ejercicio es una aplicación simple de la distribución de Poisson. Dado que se conoce la probabilidad para X=0

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

Luego

$$\lambda = -\ln(0{,}2) = 1{,}61$$

De allí

$$P(X > 2) = 1 - P(X \leq 2) = 1 - P(X = 0) - P(X = 1) - P(X = 2)$$

Reemplazando

$$P(X > 2) = 1 - 0{,}2 - \frac{\lambda^1}{1!} e^{-\lambda} - \frac{\lambda^2}{2!} e^{-\lambda} = 1 - 0{,}2 - 0{,}332 - 0{,}259 = 0{,}209$$

4. Sea X una distribución de Poisson con parámetro λ. Encontrar el valor de k para el cual $P(X = k)$ es máxima [Indicación: comparar $P(X = k)$ con $P(X = k - 1)$].

El desarrollo de la acotación de la moda fue elaborado en la teoría y nos remitimos a él. Se obtuvo que

$$\lambda - 1 \leq k_m \leq \lambda$$

Si lo aplicamos al ejercicio 2 obtenemos que el máximo de probabilidad o moda se obtiene para $k_m = 3$, el entero acotado entre 3,65-1 y 3,65. En el ejercicio 3 el valor modal se obtiene para $k_m = 1$, como puede verificarse con la probabilidad de 0,332.
En el ejercicio 1 el valor $\lambda = 2$ nos dice que la moda está entre $1 \leq k_m \leq 2$. Ambos cálculos se realizaron obteniendo $P(X = 1) = P(X = 2) = 0{,}271$. Hay una "degeneración" modal cuando dos modas tienen la misma probabilidad a la vez máxima.

5. Supóngase que la probabilidad de que un artículo producido por una máquina especial sea defectuoso vale 0,2. Si 10 artículos producidos se seleccionan al azar, ¿cuánto vale la probabilidad de que no se encuentre más de un artículo defectuoso? Use las distribuciones binomial y de Poisson, y compare las respuestas.

En este ejercicio el valor de la probabilidad del experimento binomial no es pequeño y el número de repeticiones no es grande. En principio no es válida la aproximación por Poisson. Se sugiere realizar ambos cálculos para comparar los resultados y evaluar en qué medida difieren.

$$P(X \leq 1) = P(X = 0) + P(X = 1)$$

Utilizando una binomial

$$P(X \leq 1) = \binom{10}{0} 0{,}2^0 (1 - 0{,}2)^{10-0} + \binom{10}{1} 0{,}2^1 (1 - 0{,}2)^{10-1} = 0{,}376$$

Utilizando Poisson con $\lambda = 10 * 0{,}2 = 2$

$$P(X \leq 1) = \frac{2^0}{0!} e^{-2} + \frac{2^1}{1!} e^{-2} = 3e^{-2} = 0{,}406$$

Los resultados son claramente diferentes, como era de esperarse. Redondeados a dos decimales aún difieren en un 3% pero si redondeamos el resultado con un decimal, ambos están en el orden del 40%.

6. Una compañía de seguros ha descubierto que sólo alrededor de 0,1% de la población tiene cierto tipo de accidente cada año. Si los 10.000 seleccionados fueran elegidos en forma aleatoria entre la población, ¿cuánto vale la probabilidad de que exactamente tres de estos

clientes tengan un accidente de ese tipo el próximo año? Use las distribuciones binomial y de Poisson y compare las respuestas.

En la línea del ejercicio anterior, se propone comparar el comportamiento exacto de la distribución binomial con la estimación aproximada por Poisson para un valor de probabilidad de 0,001 y un número de repeticiones de 10000, con lo cual $\lambda = 10$. Por binomial

$$P(X = 3) = \binom{10000}{3} 0{,}001^3 (1 - 0{,}001)^{10000-3} = 0{,}0075 = 0{,}75\%$$

Por Poisson

$$P(X = 3) = \frac{10^3}{3!} e^{-10} = 0{,}0076 = 0{,}76\%$$

Vemos que la diferencia se manifiesta en el cuarto decimal, con error del orden de uno en diez mil.

7. El número de buques tanques, digamos N, que llegan cada día a cierta refinería tiene una distribución de Poisson con parámetro λ=2. Las actuales instalaciones portuarias pueden despachar tres buques al día. Si más de tres buques tanques llegan en un día, los que están en exceso deben enviarse a otro puerto.
   a) En un día determinado, ¿cuánto vale la probabilidad de tener que hacer salir buques tanques?
   b) ¿En cuánto deben aumentarse las instalaciones actuales para reducir esta probabilidad a menos del 10%?
   c) ¿Cuánto vale el número esperado de buques tanques que llegan en un día?
   d) ¿Cuánto vale el número más probable de buques tanques que llegan diariamente?
   e) ¿Cuánto vale el número esperado de buques tanques atendidos diariamente?
   f) ¿Cuánto vale el número esperado de buques tanques devueltos diariamente?

Para el punto “a” evaluamos la probabilidad de que lleguen más de tres buques tanque en un día determinado, dado que sólo se puede atender hasta tres. Luego

$$P(X > 3) = 1 - P(X \leq 3) = 1 - \frac{2^0}{0!}e^{-2} - \frac{2^1}{1!}e^{-2} - \frac{2^2}{2!}e^{-2} - \frac{2^3}{3!}e^{-2}$$
$$= 1 - \left(1 + 2 + 2 + \frac{4}{3}\right)e^{-2} = 0{,}143 = 14{,}3\%$$

Si incorporamos una dársena más, la probabilidad de rechazar buques disminuye a

$$P(X > 4) = 1 - P(X \leq 4) = 0{,}143 - \frac{2^4}{4!}e^{-2} = 0{,}143 - 0{,}090 = 0{,}053 = 5{,}3\%$$

El número esperado de buques que llegan por día coincide con el parámetro de la distribución de Poisson, que vale λ=2.

El número esperado de buques atendidos no responde a una distribución de Poisson dado que sólo se pueden atender hasta tres buques, de modo que si no llega ningún barco, no se atenderá a ninguno, si llega uno, se lo atenderá, si llegan dos, se atenderá a los dos, pero si llegan tres o más, se atenderá solamente a tres. Luego si A es la variable aleatoria asociada a los buques atendidos

$$E(A) = 0 * P(X = 0) + 1 * P(X = 1) + 2 * P(X = 2) + 3 * P(X \geq 3)$$

Como

$$P(X \geq 3) = 1 - P(X < 3) = 1 - \frac{2^0}{0!}e^{-2} - \frac{2^1}{1!}e^{-2} - \frac{2^2}{2!}e^{-2} = 0{,}323$$

Resulta

$$E(A) = 0 * 0{,}135 + 1 * 0{,}271 + 2 * 0{,}271 + 3 * 0{,}323 = 1{,}77$$

La esperanza de los buques atendidos vale E(A)=1,77.
Adelantamos un resultado que formalizaremos en la unidad siguiente. La esperanza de la suma de dos variables aleatorias coincide con la suma de las esperanzas de las variables aleatorias individuales. Si llamamos D a la variable aleatoria asociada al número de buques devueltos, como X=A+D, es decir, el número que llega resulta de los atendidos más los devueltos

$$2 = E(X) = E(A + D) = E(A) + E(D) = 1{,}77 + E(D)$$

De aquí

$$E(D) = 2 - 1{,}77 = 0{,}23$$

8. Supóngase que X tiene una distribución $N(2; 0{,}16)$. Use la tabla de distribución normal para evaluar las probabilidades siguientes probabilidades:
   a) $P(X \leq 2{,}1)$ b) $P(X > 1{,}8)$ c) $P\ (1{,}8 < X \leq 2{,}1)$.

A partir de la información del ejercicio tenemos como datos $\mu = 2$ y $\sigma^2 = 0{,}16$, por lo tanto $\sigma = 0{,}4$. A partir de estos datos planteamos

$$P(X \leq 2{,}1) = P\left(z \leq \frac{2{,}1 - 2}{0{,}4}\right) = P(z \leq 0{,}25) = \phi(0{,}25) = 0{,}5987$$

El valor de la probabilidad lo obtenemos de buscar en la tabla la intersección de la fila encabezada a izquierda por “0,2” y la columna “0,05”, correspondiente al segundo decimal.
Para hallar la siguiente probabilidad, recurrimos al complemento

$$P(X > 1{,}8) = 1 - p(X \leq 1{,}8) = 1 - P\left(z \leq \frac{1{,}8 - 2}{0{,}4}\right) = 1 - P(z \leq -0{,}50) = 1 - \phi(-0{,}50)$$
$$= 1 - 0{,}3085 = 0{,}6915$$

Para hallar esta probabilidad ubicamos la fila “-0,5” y la primera columna “0,00”.
Para hallar la probabilidad en un intervalo recurrimos a la diferencia entre probabilidades acumuladas.

$$P\ (1{,}8 < X \leq 2{,}1) = P\left(z \leq \frac{2{,}1 - 2}{0{,}4}\right) - P\left(z \leq \frac{1{,}8 - 2}{0{,}4}\right) = \phi(0{,}25) - \phi(-0{,}50)$$
$$= 0{,}5987 - 0{,}3085 = 0{,}2902$$

9. Se sabe que un material se fabrica con un parámetro de calidad que responde a una distribución normal con esperanza μ = 20 y varianza σ² = 4.
   a) Calcular la probabilidad de que el número de ocurrencias sea menor que 17.
   b) Calcular la probabilidad de que el número de ocurrencias sea mayor que 21.
   c) Calcular la probabilidad de que el número de ocurrencias esté entre 17 y 21.
   d) Si sólo se puede desechar el 20% del material que menos responde al criterio de calidad empleado, ¿qué límites impondría a la variable aleatoria de manera tal que el parámetro no sea menor que tal límite? ¿Cuál sería el límite si se impone que el parámetro no sea mayor que cierto límite con esa probabilidad? ¿Cuáles serían los límites si se impone que el parámetro se encuentre fuera de ellos con 10% de probabilidad a cada lado?

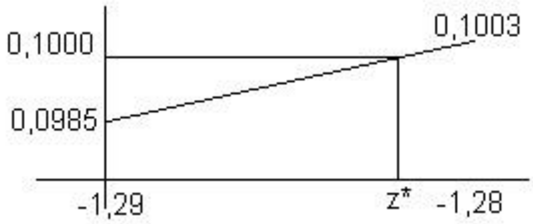
A partir de la varianza obtenemos la desviación estándar $\sigma = 2$.

$$P(X \leq 17) = P\left(z \leq \frac{17 - 20}{2}\right) = P(z \leq -1{,}50) = \phi(-1{,}50) = 0{,}0668$$

$$P(X > 21) = 1 - P(x \leq 21) = 1 - P\left(z \leq \frac{21 - 20}{2}\right) = 1 - P(z \leq 0{,}50) = 1 - \phi(0{,}50)$$
$$= 1 - 0{,}6915 = 0{,}3085$$

$$P(17 < X \leq 21) = P(X \leq 21) - P(X \leq 17) = \phi(0{,}50) - \phi(-1{,}50) = 0{,}6915 - 0{,}0668 = 0{,}6247$$

El punto "d" plantea imponer límites a la variable aleatoria de modo que quede fuera de esos límites el 20% de la probabilidad. Por simetría se dejará 10% de probabilidad en cada una de las colas de la distribución.
Si exploramos la tabla de la distribución normal, encontramos el valor de probabilidad 0,1003 correspondiente al valor $z = -1{,}28$ y, por simetría, 0,8997 a $z = 1{,}28$. El valor de probabilidad correspondiente a $z = -1{,}29$ y $z = 1{,}29$ es 0,0985 y 0,9015 respectivamente. Ninguno coincide con el 10%=0,1000 pero se puede interpolar y obtener un valor de $z = \pm 1{,}282$ como correspondiente a 10% y 90% los valores de la variable aleatoria normal estándar que dejan el 10% a cada lado de la distribución.
Esta interpolación la realizamos considerando las diferencias entre el valor de probabilidad deseado y los más próximos hallados en la tabla. Para ello observemos el gráfico que sigue



Se observa el valor de probabilidad acumulado hasta -1,29 (0,0985) y el acumulado hasta el siguiente registrado en tabla -1,28 (0,1003). Por medio de una interpolación lineal buscamos el valor de z^* que acumula el 10% de probabilidad (0,1000). Para ello buscamos la ecuación de la recta que interpola los puntos indicados. La ordenada al origen vale b=0,0985 para z=-1,29 y la pendiente

$$m = \frac{\Delta P}{\Delta z} = \frac{0{,}1003 - 0{,}0985}{-1{,}28 - (-1{,}29)} = 0{,}18 = \frac{P_s - P_i}{z_s - z_i}$$

determinan

$$P = m(z - z_i) + b = \frac{P_s - P_i}{z_s - z_i}(z - z_i) + P_i = P_i + \frac{z - z_i}{z_s - z_i}(P_s - P_i)$$

Buscamos el valor de z^* que acumule una probabilidad $P^* = 0{,}1000$.

$$0{,}1000 = P^* = P_i + \frac{z^* - z_i}{z_s - z_i}(P_s - P_i) = 0{,}0985 + \frac{z^* - (-1{,}29)}{-1{,}28 - (-1{,}29)}(0{,}1003 - 0{,}0985)$$

De allí

$$z^* = \frac{0{,}1000 - 0{,}0985}{0{,}1003 - 0{,}0985} * \big(-1{,}28 - (-1{,}29)\big) + (-1{,}29) = -1{,}281\hat{6} \cong -1{,}282$$

En general

$$z^* = z_i + \frac{P^* - P_i}{P_s - P_i} * (z_s - z_i)$$

Luego, retomando nuestro ejercicio

$$z^* = \pm 1{,}282 = \frac{L_{\pm -} - 20}{2}$$

De modo que

$$L_{\pm -} = 20 \pm 1{,}282 * 2$$

Por lo tanto los límites son $L_+ = 22{,}564$ y $L_- = 17{,}436$.
Si lo planteamos en términos de dejar fuera el 20% a cola superior, buscamos el valor de probabilidad 0,8000 en la tabla. Encontramos el valor de z=0,842 interpolando entre la probabilidad acumulada hasta 0,84 y 0,85. Luego

$$z = \pm 1{,}842 = \frac{L_{\pm -} - 20}{2}$$

De modo que

$$L_{\pm} = 20 + 0{,}842 * 2 = 21{,}684$$

Para dejar fuera el 10% de los casos, si nuevamente exploramos la tabla de la distribución normal, encontramos el valor de probabilidad 0,0505 correspondiente al valor $z = -1{,}64$ y, por simetría, 0,9495 a $z = 1{,}64$. El valor de probabilidad correspondiente a $z = -1{,}65$ y $z = 1{,}65$ es 0,0495 y 0,9505 respectivamente. Ninguno coincide con el 5%=0,0500 pero es fácil interpolar y obtener un valor de $z = \pm 1{,}645$ como correspondiente a 5% y 95% los valores de la variable aleatoria normal estándar que dejan el 5% a cada lado de la distribución. Luego

$$z = \pm 1{,}645 = \frac{L_{\pm -} - 20}{2}$$

De modo que

$$L_{\pm -} = 20 \pm 1{,}645 * 2$$

Por lo tanto los límites son $L_+ = 23{,}29$ y $L_- = 16{,}71$.

10. El diámetro de un cable eléctrico está distribuido normalmente con esperanza 0,8 y varianza 0,0004. ¿Cuánto vale la probabilidad de que el diámetro sobrepase 0,81cm? Suponiendo que el cable se considere defectuoso si el diámetro se diferencia de su esperanza en más de 0,025, ¿cuánto vale la probabilidad de obtener un cable defectuoso?

Se informa la esperanza $\mu = 0{,}8cm$ y la varianza $\sigma^2 = 0{,}0004cm^2$, de donde obtenemos $\sigma = 0{,}02cm$. Luego

$$P(D > 0{,}81cm) = 1 - P(D \leq 0{,}81cm) = 1 - P\left(z \leq \frac{0{,}81cm - 0{,}80cm}{0{,}02cm}\right) =$$
$$1 - P(z \leq 0{,}50) = 1 - \phi(0{,}50) = 1 - 0{,}6915 = 0{,}3985$$

Si se rechaza cuando la diferencia entre el valor real y la esperanza excede 0,025, los límites se obtienen sumando y restando este valor de la esperanza. Los límites valen entonces $0{,}8cm + 0{,}025cm = 0{,}825cm$ a derecha y $0{,}8cm - 0{,}025cm = 0{,}775cm$.

$$P(|X - 0{,}8cm| > 0{,}025cm) = P(X < 0{,}775cm) + P(X > 0{,}825cm) =$$

$$= P\left(z < \frac{0{,}775cm - 0{,}8cm}{0{,}02cm}\right) + 1 - P\left(z < \frac{0{,}825cm - 0{,}8cm}{0{,}02cm}\right) = 0{,}1056 + 1 - 0{,}8944$$
$$= 0{,}1056 + 0{,}1056 = 0{,}2112$$

Notemos también que formalmente debemos escribir $P(X \leq x)$ pero en la integración sobre una distribución continua es numéricamente equivalente a $P(X < x)$. También observamos que por simetría basta calcular la probabilidad a una cola y duplicarla.

11. Se sabe que los errores en cierto instrumento para medir longitudes están distribuidos normalmente con valor esperado cero y desviación estándar 1cm. ¿Cuánto vale la probabilidad de que al medir, los errores sean mayores que 1cm, 2cm y 3cm?

El objetivo del ejercicio es fijar tres valores de probabilidad de muy frecuente uso en el cálculo de probabilidades. El valor $z = \pm 1$ determina los puntos de inflexión de la densidad normal.

$$P(|X| > 1cm) = P(X < -1cm) + P(X > 1cm) = 2 * P\left(z < \frac{-1-0}{1} = -1\right) = 2 * 0{,}1587$$
$$= 0{,}3174 = 31{,}74\%$$

De modo que dentro del intervalo determinado por los puntos de inflexión de la distribución normal se encuentra una probabilidad de $0{,}6826 \cong 68\%$. En síntesis

$$P(|z| \leq 1) \cong 68\%$$

Sobre la misma base, que la variable aleatoria exceda dos desvíos estándar

$$P(|z| > 2) = 0{,}0456 \cong 5\%$$

Y con tres desvíos

$$P(|z| > 3) = 0{,}0026 \cong 0{,}25\%$$

12. Suponiendo que la duración de los instrumentos electrónicos D1 y D2 tienen distribuciones N(40;36) y N(45;9) respectivamente, ¿cuál debe preferirse para usarlo durante un período de 45 horas? ¿Cuál debe preferirse para usarlo durante un período de 48 horas?

Para evaluar cuál instrumento es conveniente para ser usado durante 45h debemos calcular la probabilidad de que dure más de ese tiempo indicado. Luego, considerando la esperanza de 40h y el desvío estándar de 6h para el primer modelo

$$P(X > 45h) = 1 - P(X \leq 45h) = 1 - P\left(z \leq \frac{45h - 40h}{6h}\right) = 1 - P(z \leq 0{,}833)$$
$$= 1 - 0{,}7986 = 0{,}2014 \cong 20\%$$

El valor de probabilidad fue interpolado entre el correspondiente a 0,83 y 0,84 en la tabla. Por otra parte, la probabilidad de que dure más de 45h con el segundo modelo vale 50% dado que coincide con la esperanza de una distribución simétrica.
En el caso de requerir que dure más de 48h tenemos que realizar los dos cálculos, teniendo en cuenta ahora que el desvío estándar de la segunda distribución vale 3h.

$$P(X > 48h) = 1 - P(X \leq 48h) = 1 - P\left(z \leq \frac{48h - 40h}{6h}\right) = 1 - P(z \leq 1{,}333)$$
$$= 1 - 0{,}9093 = 0{,}0907 \cong 9{,}1\%$$

Con el segundo modelo

$$P(X > 48h) = 1 - P(X \leq 48h) = 1 - P\left(z \leq \frac{48h - 45h}{3h}\right) = 1 - P(z \leq 1{,}00)$$
$$= 1 - 0{,}8413 = 0{,}1587 \cong 15{,}9\%$$

En ambos casos es conveniente el segundo modelo. En este ejercicio es conveniente graficar ambas distribuciones en escala para comparar las áreas correspondientes a la integración.

13. Un instrumento tiene una duración que responde a una distribución exponencial de parámetro 0,05 medido en meses. Calcular la probabilidad de que el instrumento dure un año, dos años y tres años. Indicar su esperanza y varianza.

Si se propone un valor de λ=0,05 “medido en meses”, lo que se plantea es que la unidad está expresada en la recíproca de esa unidad de tiempo, es decir $E(T) = \mu = \lambda^{-1} = 20meses$. La esperanza y la desviación estándar coinciden con la recíproca del parámetro, de modo que la varianza está dada por el cuadrado $V(T) = \sigma^2 = \lambda^{-2} = 400mes^2$.

Como se trata de un instrumento, el planteo nos lleva a que puede ser utilizado hasta que presente la primera falla, luego se trata de hallar la probabilidad de que dure más del tiempo en que se pretende darle uso. En términos de ecuaciones

$$P(t > 12meses) = \int_{12}^{\infty} \lambda e^{-\lambda t} dt = e^{-12*0,05} = 0,47$$

Para dos años

$$P(t > 24meses) = \int_{24}^{\infty} \lambda e^{-\lambda t} dt = e^{-24*0,05} = 0,30$$

Y para tres años

$$P(t > 36meses) = \int_{36}^{\infty} \lambda e^{-\lambda t} dt = e^{-36*0,05} = 0,165$$

14. Sea T la duración de un instrumento electrónico en meses con densidad:

$$f(t) = \lambda e^{-\lambda t}\ para\ t > 0$$

Si λ=0,02, ¿por cuántos meses de duración debe garantizar el fabricante sus instrumentos si busca que la probabilidad de que satisfaga la garantía valga 0,8?

Se plantea concretamente que se desea que el instrumento funcione al menos un cierto tiempo, de modo que el planteo sería

$$P(t > t^*) = \int_{t^*}^{\infty} \lambda e^{-\lambda t} dt = e^{-t^* * 0,02 meses^{-1}} = 0,8$$

Aplicando logaritmos

$$\ln(0,8) = -t^* * 0,02 meses^{-1}$$

De donde

$$t^* = -\frac{\ln(0,8)}{0,02} meses = 11,16 meses$$

15. Se asume que la duración de una conversación telefónica es una variable aleatoria X con función de densidad $f(x) = ae^{-\alpha x}$. Mostrar que la probabilidad de que una conversación dure más que $t_1 + t_2$ minutos, dado que ha durado por lo menos t_1 minutos, es igual a la probabilidad no condicional de que ésta dure por lo menos t_2 minutos. (Sugerencia: elaborar un ejemplo numérico y luego plantear el problema en general en términos de probabilidad condicional).

Tomaremos como referencia un ejemplo numérico. Supongamos que una conversación telefónica ha durado diez minutos ($t_1 = 10\text{min}$), nos cuestionamos si esta duración previa influye sobre la probabilidad de que dure otros cinco minutos más ($t_2 = 5\text{min}$), con lo cual debería durar al menos ($t_1 + t_2 = 15\text{min}$). Formalmente el problema planteado es

$$P(t > t_1 + t_2/t > t_1) = \frac{P\big(t \in [t_1 + t_2; \infty) \cap [t_1; \infty)\big)}{P\big(t \in [t_1; \infty)\big)} = \frac{P\big(t \in [t_1 + t_2; \infty)\big)}{P\big(t \in [t_1; \infty)\big)}$$

Desarrollando

$$P(t > t_1 + t_2/t > t_1) = \frac{e^{t_1 + t_2}}{e^{t_1}} = e^{t_2} = P(t > t_2)$$

Interpretamos el resultado final considerando que la probabilidad de que la llamada dure más de quince minutos sabiendo que ya duró diez minutos, es la misma que la probabilidad de que dure cinco minutos independientemente de la duración previa. A esta propiedad de la distribución exponencial la habíamos llamado de “falta de memoria”.

16. Compare la *cota superior* de la probabilidad $P[|X-E(X)|>2\sqrt{V(X)}]$ obtenida con la desigualdad de Tchebychev con la probabilidad exacta en cada uno de los casos siguientes:
    a) $P\big(|X-E(X)|>2\sigma(X)\big)$ si X tiene distribución N(μ;σ²) y
    b) $P\left(|X-E(X)|>\frac{2}{\lambda}\right)$ si X tiene distribución exponencial con parámetro λ.

Si recordamos que el teorema de Tchebyshev expresa que $P(|X-\mu|>k\sigma)<\frac{1}{k^2}$, en nuestro caso

$$P(|X-\mu|>2\sigma)<\frac{1}{2^2}=\frac{1}{4}=0{,}25=25\%$$

La aplicación específica a una distribución normal y a una distribución exponencial nos lleva a obtener

$$P(|X-\mu|>2\sigma)=2\phi(-2)=2*0{,}0228=0{,}056=5{,}6\%<25\%$$

Lo cual verifica la desigualdad de Tchebyshev en el caso particular de una normal. Si es una distribución exponencial

$$P(|X-\lambda^{-1}|>2\lambda^{-1})=P(x-\lambda^{-1}>2\lambda^{-1})=P(x>3\lambda^{-1})=e^{-3\lambda\lambda^{-1}}=e^{-3}=0{,}0498 \\ \cong 5\%<25\%$$

Lo que también verifica la desigualdad de Tchecyshev para una exponencial.

17. La vida media de un átomo radioactivo se define como el tiempo que tarda en desintegrarse la mitad del material radioactivo. El proceso de desintegración responde a una distribución exponencial de parámetro λ ($N=N_0e^{-\lambda t}$). Si se mide el tiempo de vida media y resulta ser de 10 años, hallar el valor de λ. ¿A qué parámetro estadístico asociaría el tiempo de vida medio? Comparar el tiempo de vida medio con la esperanza de la exponencial, interpretar esta diferencia en relación con la forma de la distribución. Hallar la probabilidad de que en un año, un mes, un día y una hora se registre una desintegración en un átomo determinado. Si la muestra contiene 10000 átomos, hallar la probabilidad de que en una hora se registren dos desintegraciones.

Para hallar el valor de λ consideramos que $t_{1/2}=10años$, luego

$$N=\frac{N_0}{2}=N_0e^{-\lambda t_{1/2}}$$

Utilizando logaritmos

$$-\ln(2)=-\lambda t_{\frac{1}{2}}=-\lambda*10años$$

Por lo tanto

$$\lambda=\frac{\ln(2)}{10años}=0{,}0693año^{-1}$$

El tiempo de vida medio lo asociamos con la esperanza considerando que $\mu=\lambda^{-1}=14{,}4años$, y éste es el tiempo que tarda en decaer a la proporción dada por e^{-1} con respecto al número inicial de átomos, y el tiempo de vida medio es el que tarda en decaer a 2^{-1} con respecto al estado inicial, delimitando dos puntos en el proceso de decaimiento expresado por la función de distribución, uno coincidente con la mediana y otro con el tiempo esperado.

Considerando la expresión del proceso Poisson, primero consideremos un tiempo de diez años y la ocurrencia de ningún evento, con lo cual

$$P(X=k/T=t)=\frac{(\lambda t)^k}{k\,!}e^{-\lambda t}=\frac{\left(\frac{\ln(2)}{10años}10años\right)^0}{0!}e^{-\frac{\ln(2)}{10años}10años}=0{,}5=50\%$$

Como era de esperarse. Calculamos ahora la probabilidad de ocurrencia de un evento en diez años.

$$P(X=k/T=t)=\frac{(\lambda t)^k}{k\,!}e^{-\lambda t}=\frac{\left(\frac{\ln(2)}{10años}10años\right)^1}{1!}e^{-\frac{\ln(2)}{10años}10años}=0{,}5*\ln(2)=0{,}35$$
$$=35\%$$

Para que se registre una desintegración en un año

$$P(X=k/T=t)=\frac{(\lambda t)^k}{k\,!}e^{-\lambda t}=\frac{\left(\frac{\ln(2)}{10años}1años\right)^1}{1!}e^{-\frac{\ln(2)}{10años}1años}=6{,}5\%$$

Para que se registre una desintegración en un mes

$$P(X=k/T=t)=\frac{(\lambda t)^k}{k\,!}e^{-\lambda t}=\frac{\left(\frac{\ln(2)}{10años}\frac{1}{12}años\right)^1}{1!}e^{-\frac{\ln(2)}{10años}\frac{1}{12}años}=0{,}56\%$$

Para que se registre una desintegración en un día

$$P(X=k/T=t)=\frac{(\lambda t)^k}{k\,!}e^{-\lambda t}=\frac{\left(\frac{\ln(2)}{10años}\frac{1}{365}años\right)^1}{1!}e^{-\frac{\ln(2)}{10años}\frac{1}{365}años}=0{,}019\%$$

Para que se registre una desintegración en una hora

$$P(X=k/T=t)=\frac{(\lambda t)^k}{k\,!}e^{-\lambda t}=\frac{\left(\frac{\ln(2)}{10años}\frac{1}{8760}años\right)^1}{1!}e^{-\frac{\ln(2)}{10años}\frac{1}{8760}años}=0{,}0008\%$$

Unidad 6: Distribuciones multivariadas

Distribuciones multivariadas

Si bien hasta el momento se ha trabajado con distribuciones sobre una única variable aleatoria, la realidad muestra siempre un patrón complejo de variabilidad en el que intervienen más de una fuente de variación. Para abordar el problema se tratará el caso de dos variables aleatorias simultáneas o “conjuntas” $(X1; X2)$. Implícitamente este problema ha sido tratado en la unidad 2 cuando se planteó la ocurrencia de dos eventos A y B en términos de una intersección $P(A \cap B)$. Expresado en la terminología de las variables aleatorias resulta

$$P(X1 = x_1; X2 = x_2) = p_{12}$$

A este valor se le llamará “probabilidad conjunta” de *X1* y *X2* como probabilidad de ocurrencia de los dos eventos.

La probabilidad asociada a una variable ignorando el comportamiento de la otra, es decir, sea cual fuere el valor de la otra, se llamará “probabilidad marginal” y notará $P_1(X1 = x_1)$ o bien $P_2(X2 = x_2)$, mientras que la probabilidad de un evento sobre una variable si se dispone de la información acerca de la ocurrencia de un evento en la otra variable se llamará “probabilidad condicional” y escribirá

$$P(X1 = x_1/X2 = x_2) = p_{1|2}$$

En primer lugar observamos que las variables aleatorias multivariadas pueden ser discretas o continuas. En el caso de las distribuciones discretas, la distribución de probabilidad se expresa en términos de una tabla de doble entrada. Las distribuciones marginales se obtienen

$$P_1(X1 = x_{1i}) = \sum_{j=1}^{n} P\left(X1 = x_{1i}; X2 = x_{2j}\right)$$

para la distribución marginal de *X1*, y

$$P_2\left(X2 = x_{2j}\right) = \sum_{i=1}^{m} P\left(X1 = x_{1i}; X2 = x_{2j}\right)$$

para la distribución marginal de *X2*. Vemos que en cada caso sumamos las probabilidades sobre el índice de la otra variable, o de las otras si hubiere más.

Por ejemplo, en la siguiente tabla

X1\X2	1	2	3	Marginal de X1
1	0,10	0.15	0.20	0.45
2	0.05	0.10	0.10	0.25
3	0.05	0.15	0.10	0.30
Marginal de X2	0.20	0.40	0.40	1.00

en filas se presenta la variable *X1* y en columnas la variable *X2* con tres valores posibles para cada una. Puede verse que P(X1=3; X2=2) = 0,15. En la fila inferior se presenta la probabilidad marginal de *X2* sumando sobre todos los valores posibles de la variable *X1*. Por otra parte, en la última columna se muestra la probabilidad marginal de *X1* sumando sobre todos los valores posibles de la variable *X2*. En el cuadro inferior derecho se muestra que la tabla efectivamente representa una distribución de probabilidad porque la suma total de todas las probabilidades conjuntas y las marginales da por resultado 1.00. Gráficamente puede presentarse por medio de un gráfico tridimensional de barras.

En el caso de distribuciones continuas en dos variables, se representan gráficamente por medio de una superficie en tres dimensiones. En tal caso se opera con funciones e integrales en dos variables. Será $f(x_1;x_2)$ la densidad conjunta y $F(x_1;x_2)$ la distribución conjunta. Se calculará

$$f_{x1}(x_1) = \int_{-\infty}^{+\infty} f(x_1;x_2)dx_2$$

para la marginal de *X1* y

$$f_{x2}(x_2) = \int_{-\infty}^{+\infty} f(x_1;x_2)dx_1$$

para la marginal de X2. Estas integrales se evalúan sobre todo el rango de la otra variable, es en el intervalo (-∞, ∞).

La distribución se interpretará

$$F(x_1;x_2) = P(X1 \leq x_1; X2 \leq x_2)$$

y las probabilidades se calcularán como el área encerrada debajo de la superficie dada por $f(x_1;x_2)$ delimitada por los límites correspondientes de las variables y aplicando las herramientas del cálculo diferencial e integral multivariado.

Distribución condicional

Recordando la definición de probabilidad condicional y de función de distribución, notaremos, si la variable aleatoria es discreta

$$F_1(x_{1k}/X_2 = x_2) = \frac{\sum_{i=1}^{k} P(X_1 = x_{1i}; X_2 = x_2)}{P(X_2 = x_2)}$$

Si es continua

$$F_1(x_1/X_2 = x_2) = \frac{\int_{-\infty}^{x_1} f(x_1;x_2)dx_1}{P(X_2 = x_2)}$$

Esta última expresión carece de sentido si la variable aleatoria X_2 es continua. En el caso de una distribución continua en ambas variables expresamos la función de distribución condicional

$$F(x_1/X_2 \leq x_2) = P(X_1 \leq x_1/X_2 \leq x_2) = \frac{P(X_1 \leq x_1; X_2 \leq x_2)}{P(X_2 \leq x_2)} = \frac{\int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f(x_1;x_2)dx_1dx_2}{\int_{-\infty}^{x_2} f_2(x_2)dx_2}$$
$$= \int_{-\infty}^{x_1} f_{x_1/x_2}(x_1)dx_1$$

La última expresión sólo puede ser válida para todo par de valores x_1 y x_2 si definimos la notación dentro del último símbolo de integración como función de densidad condicional de la variable aleatoria X1 condicionada por la variable aleatoria X2 como

$$f(x_1/X_2 = x_2) = f_{x_1/x_2}(x_1) = \frac{f(x_1;x_2)}{f_2(x_2)}$$

Independencia entre variables aleatorias

Si recordamos que cuando dos eventos son independientes la probabilidad conjunta resulta $P(A \cap B) = P(A).P(B)$, y notamos que $P(X1 \leq x_1; X2 \leq x_2)$ es una expresión de dos eventos en términos de sus variables aleatorias asociadas, si éstas son independientes obtenemos la expresión $P(X1 \leq x_1; X2 \leq x_2) = P(X1 \leq x_1).P(X2 \leq x_2)$. Desarrollado en términos de las funciones de distribución

$$F(x_1; x_2) = F(x_1).F(x_2)$$

Escrito en términos de las integrales de las funciones de densidad respectivas y conjunta

$$\iint_{-\infty}^{x_1;x_2} f(x_1; x_2)dx_1dx_2 = \int_{-\infty}^{x_1} f(x_1)dx_1 \,. \int_{-\infty}^{x_2} f(x_2)dx_2$$

Como esta relación vale para todo par $(x_1; x_2)$, los integrandos tienen que ser iguales, de modo que, si dos variables $(x_1; x_2)$ son independientes, expresando la última igualdad en el formato de las densidades marginales

$$f(x_1; x_2) = f(x_1).f(x_2) = f_1(x_1).f_2(x_2)$$

Esperanza y varianza de una suma de variables aleatorias. Covarianza

Uno de los problemas que se debe abordar es el de las combinaciones lineales y transformaciones lineales de variables aleatorias. Comenzaremos por trabajar sobre dos variables aleatorias por medio de la esperanza de la suma.

Sean X e Y dos variables aleatorias con función de densidad conjunta $f(x; y)$. La esperanza de la suma de las variables aleatorias se obtiene

$$\begin{aligned} E(X+Y) &= \iint_{-\infty}^{+\infty} (x+y)f(x;y)dxdy \\ &= \int_{-\infty}^{+\infty} x \int_{-\infty}^{+\infty} f(x,y)\, dy\, dx + \int_{-\infty}^{+\infty} y \int_{-\infty}^{+\infty} f(x,y)\, dx\, dy \\ &= \int_{-\infty}^{+\infty} x f_x(x)dx + \int_{-\infty}^{+\infty} y f_y(y)dy = E(X) + E(Y) \end{aligned}$$

Para la varianza en cambio nos veremos obligados a introducir un nuevo parámetro que tendrá importancia central

$$\begin{aligned} V(X+Y) &= \iint_{-\infty}^{+\infty} \big((x+y) - E(X+Y)\big)^2 f(x;y)dxdy \\ &= \iint_{-\infty}^{+\infty} \big(x - E(X)\big)^2 + 2\big(x - E(X)\big)\big(y - E(Y)\big) + \big(y - E(Y)\big)^2 f(x;y)dxdy \\ &= V(X) + V(Y) + 2C(X,Y) \end{aligned}$$

donde

$$Cov(X;Y) = \iint_{-\infty}^{+\infty} \big(x - E(X)\big)\big(y - E(Y)\big)f(x;y)dxdy$$

se define como la “covarianza” entre las variables aleatorias X e Y. Probaremos que, si las dos variables son independientes, la integral es separable y la covarianza es nula

$$Cov(X;Y) = \int_{-\infty}^{+\infty} \big(x - E(X)\big)f_x(x)dx \int_{-\infty}^{+\infty} \big(y - E(Y)\big)f_y(y)dy$$

$$Cov(X;Y) = \left(\int_{-\infty}^{+\infty} x f_x(x)dx - E(X)\int_{-\infty}^{+\infty} f_x(x)dx\right)\left(\int_{-\infty}^{+\infty} y f_y(y)dy - E(Y)\int_{-\infty}^{+\infty} f_y(y)dy\right)$$
$$= (E(X) - E(X).1)(E(Y) - E(Y).1) = 0$$

Por lo tanto, si las variables aleatorias son independientes, la varianza de la suma de variables aleatorias es la suma de las varianzas $V(X+Y) = V(X) + V(Y)$. Si las variables fuesen discretas, las integrales son reemplazadas por sumatorias y se obtiene el mismo resultado.

Esperanza condicional

Sobre la base de la noción de esperanza y de probabilidad condicional, dada un par de variables aleatorias $(X_1;X_2) = \{(x_{1i};x_{2j})\}$ definimos la esperanza condicional

$$E(X_1/X_2 = x_2) = \sum_{j=1}^{n} x_{1i} * p(x_1/x_{2j})$$

$$E(X_2/X_1 = x_1) = \sum_{i=1}^{m} x_{2j} * p(x_2/x_{1i})$$

Para cada valor de $X_2 = \{x_{2j}; j = 1 \dots n\}$ habrá un valor diferente de la esperanza de X_1, y recíprocamente, para cada valor de $X_1 = \{x_{1i}; i = 1 \dots m\}$ habrá un valor de esperanza para X_2.

A partir de la definición de función de densidad condicional, tenemos para el caso continuo

$$E(X_1/X_2 = x_2) = \int_{-\infty}^{+\infty} x_1 f_{x_1/x_2}(x_1)dx_1 = \int_{-\infty}^{+\infty} x_1 \frac{f(x_1;x_2)}{f_2(x_2)} dx_1$$

$$E(X_2/X_1 = x_1) = \int_{-\infty}^{+\infty} x_2 f_{x_2/x_1}(x_2)dx_2 = \int_{-\infty}^{+\infty} x_2 \frac{f(x_1;x_2)}{f_1(x_1)} dx_2$$

Probabilidades y Estadística: Ejercicios Unidad Nº 6

1. Suponga que la tabla siguiente representa la distribución de probabilidades conjunta de la variable aleatoria discreta (X , Y). Completar la tabla. Calcular todas las distribuciones marginales y condicionales:

X1\X2	1	2	3	
1	0.10	0.15	0.10	
2	0.05	0.10	0.05	
3	0.15	0.15	¿?	

2. Supóngase que la función de densidad de probabilidad conjunta de (X , Y) está dada por:

$$f(x,y) = \begin{cases} e^{-(x+y)} & x \geq 0 \quad y \geq x \\ 0 & x < 0 \quad o \quad y < 0 \end{cases}$$

a) Encontrar la función de densidad de probabilidad marginal de X.
b) Encontrar la función de densidad de probabilidad marginal de Y.
c) Evaluar P(X > 2 | Y < 4).

Resolución Ejercicios Unidad Nº 6

1. Suponga que la tabla siguiente representa la distribución de probabilidades conjunta de la variable aleatoria discreta (X;Y). Completar la tabla. Calcular todas las distribuciones marginales y condicionales:

X1\X2	1	2	3	
1	0.10	0.15	0.10	0.35
2	0.05	0.10	0.05	0.20
3	0.15	0.15	0.15	0.45
	0.30	0.40	0.30	1.00

Considerando que la suma de todas las probabilidades debe dar "1", completamos la tabla con el dato faltante. Calculamos todas las marginales. A modo de ejemplo, la probabilidad de que X1=2 sabiendo que X2=3 debería plantearse y calcularse…

$$P(X1 = 2/X2 = 3) = \frac{P(X1 = 2 \wedge X2 = 3)}{P(X2 = 3)} = \frac{0,05}{0,30} = \frac{1}{6} = 0,166 \ldots$$

2. Supóngase que la función de densidad de probabilidad conjunta de (X , Y) está dada por:

$$f(x,y) = \begin{cases} e^{-(x+y)} & x \geq 0 \quad y \geq x \\ 0 & x < 0 \quad o \quad y < 0 \end{cases}$$

a) Encontrar la función de densidad de probabilidad marginal de X.
b) Encontrar la función de densidad de probabilidad marginal de Y.
c) Evaluar P(X > 2 | Y < 4).

Para calcular las marginales integramos sobre el rango de la otra variable. Primero calcularemos la marginal de la variable aleatoria "x". Para ello fijamos el dominio como un triángulo infinito en la parte superior del primer cuadrante acotado por el eje vertical "y" y por la recta $y = x$.

$$f_x(x) = \int_{-\infty}^{+\infty} f(x;y)dy = \int_{x}^{+\infty} e^{-(x+y)}dy = e^{-x}\int_{x}^{+\infty} e^{-y}dy = -e^{-x}e^{-y}\Big|_{x}^{+\infty}$$
$$= -e^{-x}(0 - e^{-x}) = e^{-2x}$$

Para calcular la marginal de "y" integramos "x" entre cero y $x = y$.

$$f_y(y) = \int_{-\infty}^{+\infty} f(x;y)dx = \int_{0}^{y} e^{-(x+y)}dx = e^{-y}\int_{0}^{y} e^{-x}dx = -e^{-y}e^{-x}\Big|_{0}^{y}$$
$$= -e^{-y}(e^{-y} - 1) = e^{-y}(1 - e^{-y})$$

Par hallar la distribución condicional hallamos el dominio de integración para la probabilidad conjunta y calculamos la probabilidad relativa al dominio condicionante.

$$P(X > 2/Y < 4) = \frac{P(X > 2/Y < 4)}{P(Y < 4)} = \frac{\iint_{x=2;y=2}^{x=y;y=4} e^{-(x+y)}dxdy}{\int_0^4 f_y(y)dy}$$

$$= \frac{\int_2^4\left(\int_2^y e^{-(x+y)}dx\right)dy}{\int_0^4 e^{-y}(1-e^{-y})dy} = \frac{\int_2^4 e^{-y}\left(\int_2^y e^{-x}dx\right)dy}{\int_0^4(e^{-y}-e^{-2y})dy} =$$

$$= \frac{\int_2^4 e^{-y}(e^{-2}-e^{-y})dy}{\int_0^4(e^{-y}-e^{-2y})dy} = \frac{(-e^{-2}e^{-y}|_2^4) - \left(-\frac{1}{2}e^{-2y}|_2^4\right)}{(-e^{-y}|_0^4) - \left(-\frac{1}{2}e^{-2y}|_0^4\right)}$$

$$= \frac{e^{-2}(e^{-2}-e^{-4}) + \frac{1}{2}(e^{-4}-e^{-8})}{(1-e^{-4}) - \frac{1}{2}(1-e^{-8})} = \frac{0{,}0158 + 0{,}0090}{0{,}4819} = 0{,}0515$$