Formula de la distribucion normal

La distribución normal o distribución gaussiana es la distribución de probabilidad en variable continua, en la que la función densidad de probabilidad está descrita por una función exponencial de argumento cuadrático y negativo, que da lugar a una forma acampanada. El nombre de distribución normal viene del hecho que esta distribución es la que se aplica a mayor número de situaciones donde está involucrada alguna variable aleatoria continua en un grupo o población dada. Figura 1. Distribución normal $N(x; \mu,\sigma)$ y su densidad de probabilidad $f(s; \mu,\sigma)$. (Elaboración propia) Como ejemplos donde se aplica la distribución normal se tienen: la altura de los hombres o de las mujeres, variaciones en la medida de alguna magnitud física o en rasgos psicológicos o sociológicos medibles como el cociente intelectual o los hábitos de consumo de cierto producto. Por otra parte, se le llama distribución gaussiana o campana de Gauss, porque es a este genio matemático alemán a quién se le acredita su descubrimiento por el uso que le dio para la descripción del error estadístico de las mediciones astronómicas allá por el año 1800. Sin embargo, se afirma que esta distribución estadística fue publicada previamente por otro gran matemático de origen francés, como lo fue Abraham de Moivre, allá por el año 1733. [toc] Fórmula A la función distribución normal en la variable continua x, con parámetros μ y σ se le denota por: $N(x; \mu,\sigma)$ y explícitamente se escribe así: $N(x; \mu,\sigma) = \int_{-\infty}^{x} f(s; \mu,\sigma)\, ds$ donde $f(u; \mu,\sigma)$ es la función densidad de probabilidad: $f(s; \mu,\sigma) = (1/(\sigma\sqrt{2\pi})) \operatorname{Exp}(-s^2/(2\sigma^2))$ La constante que multiplica a la función exponencial en la función densidad de probabilidad se le llama constante de normalización, y se ha elegido de tal manera que: $N(+\infty, \mu,\sigma) = 1$ La expresión anterior asegura que la probabilidad de que la variable aleatoria x esté comprendida entre $-\infty$ y $+\infty$ sea 1, es decir el 100% de probabilidad. El parámetro μ es la media aritmética de la variable aleatoria continua x y σ la desviación típica o raíz cuadrada de la varianza de esa misma variable. En el caso que $\mu = 0$ y $\sigma = 1$ se tiene entonces la distribución normal estándar o distribución normal típica: $N(x; \mu = 0, \sigma = 1)$ Características de la distribucion normal 1- Si una variable estadística aleatoria sigue una distribución normal de densidad de probabilidad $f(s; \mu,\sigma)$, la mayor parte de los datos se agrupan alrededor de valor medio μ y están dispersos a su alrededor de forma tal que poco más de ⅔ de los datos están entre $\mu - \sigma$ y $\mu + \sigma$. Puede servirte: Prisma heptagonal2- La desviación típica σ siempre es positiva. 3- La forma de la función de densidad f se asemeja a la de una campana, por lo que a esta función muchas veces se le llama campana de Gauss o función gaussiana. 4- En una distribución gaussiana la media, la mediana y la moda coinciden. 5- Los puntos de inflexión de la función densidad de probabilidad se encuentran justamente en $\mu - \sigma$ y $\mu + \sigma$. 6- La función f es simétrica respecto a un eje que pase por su valor medio μ y tiene asintóticamente a cero para $x \longrightarrow +\infty$ y $x \longrightarrow -\infty$. 7- A mayor valor de σ mayor dispersión, ruido o distanciamiento de los datos alrededor del valor medio. Es decir a mayor σ la forma de campana es más abierta. En cambio σ pequeño indica que los dados se ciñen a la media y la forma de la campana es más cerrada o puntiaguda. 8- La función de distribución $N(x; \mu,\sigma)$ indica la probabilidad que la variable aleatoria sea menor o igual que x. Por ejemplo, en la figura 1 (más arriba) la probabilidad P de que la variable x sea menor o igual a 1.5 es de 84% y se corresponde con el área bajo la función densidad de probabilidad $f(x; \mu,\sigma)$ desde $-\infty$ hasta x. Intervalos de confianza 9- Si los datos siguen una distribución normal, entonces 68,26% de estos están entre $\mu - \sigma$ y $\mu + \sigma$. 10- El 95,44% de los datos que siguen una distribución normal se encuentran entre $\mu - 2\sigma$ y $\mu + 2\sigma$. 11- El 99,74% de los datos que siguen una distribución normal se encuentran entre $\mu - 3\sigma$ y $\mu + 3\sigma$. 12- Si una variable aleatoria x sigue una distribución $N(x; \mu,\sigma)$, entonces la variable $z = (x - \mu) / \sigma$ sigue la distribución normal estándar $N(z; 0,1)$. El cambio de la variable x a la z recibe el nombre de estandarización o tipificación y es de gran utilidad a al momento de aplicar las tablas de la distribución estándar a los datos que siguen una distribución normal no-estándar. Aplicaciones de la distribución normal Para aplicar la distribución normal es necesario pasar por el cálculo de la integral de la densidad de probabilidad, lo cual desde el punto de vista analítico no es fácil y no siempre se dispone de un programa informático que permita su cálculo numérico. Para este fin se usan las tablas de valores normalizados o tipificados, que no es más que la distribución normal en el caso $\mu = 0$ y $\sigma = 1$. Puede servirte: Tipos de ángulos, características y ejemplosTabla de distribución normal tipificada (parte 1/2) Tabla de distribución normal tipificada (parte 2/2) Debe notarse que estas tablas no incluyen los valores negativos. Sin embargo, usando las propiedades de simetría de la función densidad de probabilidad gaussiana pueden obtenerse los valores correspondientes. En el ejercicio resuelto mostrado más abajo se indica el uso de la tabla en estos casos. Ejemplo Suponga que tiene un conjunto de datos aleatorios x que siguen una distribución normal de media 10 y desviación típica 2. Se pide encontrar la probabilidad de que: a) La variable aleatoria x sea menor o igual a 8. b) Sea menor o igual a 10. c) Que la variable x esté por debajo de 12. d) La probabilidad que un valor x esté entre 8 y 12. Solución: a) Para responder a la primera pregunta simplemente hay que calcular: $N(x; \mu,\sigma)$ Con $x = 8$, $\mu = 10$ y $\sigma = 2$. Nos percatamos que se trata de una integral que no tiene una solución analítica en funciones elementales, sino la solución está expresada en función de la función error erf(x). Por otra parte, existe la posibilidad de resolver la integral en forma numérica, que es lo que hacen muchas calculadoras, las hojas de cálculo y programas informáticos como GeoGebra. La siguiente figura muestra la solución numérica correspondiente al primer caso: Figura 2. Densidad de probabilidad $f(x; \mu,\sigma)$. El área sombreada representa $P(x \leq 8)$. (Elaboración propia) y la respuesta es que la probabilidad que x esté por debajo de 8 es: $P(x \leq 8) = N(x=8; \mu=10,\sigma=2) = 0{,}1587$ b) En este caso se trata de encontrar la probabilidad de que la variable aleatoria x esté por debajo de la media que en este caso vale 10. La respuesta no requiere cálculo alguno, ya que sabemos que la mitad de los datos están por debajo de la media y la otra mitad por encima de la media. Por ello, la respuesta es: $P(x \leq 10) = N(x=10; \mu=10,\sigma=2) = 0{,}5$ c) Para responder a esta pregunta hay que calcular $N(x=12; \mu=10,\sigma=2)$, lo cual puede hacerse con una calculadora que tenga funciones estadísticas o mediante un software como es el caso de GeoGebra: Puede servirte: Eventos independientes: demostración, ejemplos, ejerciciosFigura 3. Densidad de probabilidad $f(x; \mu,\sigma)$. El área sombreada representa $P(x \leq 12)$. (Elaboración propia) La respuesta a la parte c puede verse en la figura 3 y es: $P(x \leq 12) = N(x=12; \mu=10,\sigma=2) = 0{,}8413$. d) Para encontrar la probabilidad de que la variable aleatoria x esté comprendida entre 8 y 12 podemos usar los resultados de las partes a y c de la siguiente manera: $P(8 \leq x \leq 12) = P(x \leq 12) - P(x \leq 8) = 0{,}8413 - 0{,}1587 = 0{,}6826 = 68{,}26\ \%$. Ejercicio resuelto El precio promedio de las acciones de una empresa es de \$25 con una desviación típica de \$4. Determine la probabilidad que: a) Una acción tenga un costo menor de \$20. b) Que tenga un costo mayor de \$30. c) El precio esté comprendido entre \$20 y \$30. Usar las tablas de distribución normal tipificada para encontrar las respuestas. Solución: Para poder hacer uso de las tablas, es necesario pasar a la variable z normalizada o tipificada: \$20 en la variable normalizada equivale a z = (\$20 – \$25) / \$4 = -5/4 = -1,25 y \$30 en la variable normalizada equivale a z = (\$30 – \$25) / \$4 = +5/4 = +1,25. a) \$20 equivale a -1,25 en la variable normalizada, pero la tabla no tiene valores negativos, por lo que ubicamos el valor +1,25 que arroja el valor de 0,8944. Si a este valor se le resta 0,5 el resultado será el área entre 0 y 1,25 que, por cierto, es idéntica (por simetría) al área entre -1.25 y 0. El resultado de la resta es 0,8944 – 0,5 = 0,3944 que es el área entre -1.25 y 0. Pero interesa el área desde $-\infty$ hasta -1,25 que será 0,5 – 0,3944 = 0,1056. Se concluye por tanto que la probabilidad que una acción esté por debajo de \$20 es 10,56%. b) \$30 en la variable tipificada z es 1,25. Para este valor en la tabla aparece el número 0,8944 que corresponde al área desde $-\infty$ hasta +1,25. El área entre +1.25 y $+\infty$ es (1 – 0,8944) = 0,1056. Es decir que la probabilidad que una acción cueste más de \$30 es 10,56%. c) La probabilidad de que una acción tenga un costo comprendido entre \$20 y \$30 se calculará así: 100% -10,56% – 10,56% = 78,88% Referencias Estadística y probabilidad. Distribución normal. Recuperado de: proyectodescartes.org Geogebra. Geogebra clásico, cálculo de probabilidad. Recuperado de geogebra.org MathWorks. Distribución de Gauss. Recuperado de: es.mathworks.com Mendenhall, W. 1981. Estadística para Administración y Economía. 3ra. edición. Grupo Editorial Iberoamérica. Stat Trek. Teach yourself Statistics. Poisson Distribution. Recuperado de: stattrek.com, Triola, M. 2012. Elementary Statistics. 11th. Ed. Pearson Education. Universidad de Vigo. Principales distribuciones continuas. Recuperado de: anapg.webs.uvigo.es Wikipedia. Distribución normal. Recuperado de: es.wikipedia.org En este artículo se explica qué es la distribución normal en estadística. Así pues, encontrarás la definición de la distribución normal, ejemplos de distribuciones normales y cuáles son las propiedades de la distribución normal.La distribución normal es una distribución de probabilidad continua cuya gráfica tiene forma de campana y es simétrica respecto a su media. En estadística, la distribución normal sirve para modelizar fenómenos de características muy diferentes, por eso es tan importante esta distribución.De hecho, en estadística la distribución normal se considera, por mucho, la distribución más importante de todas las distribuciones de probabilidad, ya que no solo permite modelizar un gran número de fenómenos reales, sino que además la distribución normal se puede usar para aproximar otros tipos de distribuciones bajo ciertas condiciones.El símbolo de la distribución normal es la letra mayúscula N. Así pues, para indicar que una variable sigue una distribución normal se indica con la letra N y se añade entre paréntesis los valores de su media aritmética y su desviación estándar.La distribución normal recibe muchos nombres diferentes, entre ellos destacan distribución de Gauss, distribución gaussiana y distribución de Laplace-Gauss.Normalmente, los conjuntos de datos que siguen una distribución normal tienen un gran número de observaciones y tratan de temas muy generales. A continuación puedes ver varios ejemplos de muestras estadísticas que típicamente se pueden modelar con una distribución normal.Ejemplos de la distribución de normal:La estatura de los alumnos de un curso.El coeficiente intelectual de los trabajadores de una empresa.El número de piezas defectuosas producidas en una fábrica durante un día.Las notas obtenidas en un examen por los alumnos de un curso.La rentabilidad de las acciones de las empresas que cotizan en bolsa.Una vez hemos visto en qué consiste la distribución normal y varios ejemplos de este tipo de distribución de probabilidad, vamos a ver cómo es su gráfica para entender mejor el concepto.En el siguiente gráfico puedes ver cómo varia la función de densidad de la distribución normal dependiendo de los valores de su media aritmética y de su desviación típica.Al tener forma de campana centrada en la media aritmética, si una variable tiene una distribución normal significa que el valor más repetido es la media y que los valores alrededor de la media se repiten con más frecuencia que los valores de los extremos. Asimismo, cuanto mayor sea la desviación típica de la distribución normal, más aplastada es la forma de su representación gráfica.Por otro lado, la gráfica de la función de probabilidad acumulada de la distribución normal también depende de los valores de su media aritmética y su desviación típica, tal y como puedes ver en la siguiente imagen:La función de densidad y la función de distribución de la distribución normal permiten calcular probabilidades relacionadas con esta distribución. No obstante, en lugar de utilizar sus fórmulas, puedes usar directamente las tablas de la distribución normal ya que es más rápido. Puedes ver estas tablas en el siguiente enlace:➤ Ver: Tabla de la distribución normalLa distribución normal tiene las siguientes características:La distribución normal depende de dos parámetros característicos que son su media aritmética (μ) y su desviación típica (σ).La distribución normal puede tomar tanto valores positivos como negativos, por lo tanto, el dominio de la distribución normal son todos los números reales.La mediana y la moda de la distribución normal son iguales a la media aritmética de la distribución.El coeficiente de asimetría y el coeficiente de curtosis de la distribución normal son nulos.La fórmula de la función de densidad de la distribución normal es la siguiente:Asimismo, la fórmula de la función de probabilidad acumulada de la distribución normal es la siguiente:Una aplicación del teorema del límite central es que una distribución de Poisson se puede aproximar a una distribución normal cuando el valor de λ es suficientemente grande.Otra aplicación del teorema del límite central es que una distribución binomial se puede aproximar a una distribución normal para conjuntos de datos con un gran número de observaciones.La distribución normal estándar, también llamada distribución normal unitaria, es el caso más simple de una distribución normal. En concreto, la distribución normal estándar es una distribución normal con valores de media y desviación estándar iguales a 0 y 1 respectivamente.Cabe destacar que cualquier distribución normal se puede transformar en una distribución normal estándar aplicando un proceso llamado tipificación que consiste en restar a cada uno de los valores su media aritmética y después dividir por su desviación típica.Además, la distribución normal estándar se usa para determinar cualquier probabilidad de cualquier distribución normal mediante su tabla de probabilidades. De manera que para hallar una probabilidad de una distribución normal primero se tipifica la variable para convertirla en una distribución normal estándar y, posteriormente, se mira en la tabla cuál es el valor de la probabilidad correspondiente. Para saber más al respecto, haz clic en el siguiente enlace:➤ Ver: Distribución normal estándarEn estadística, la regla empírica, también llamada regla 68-95-99,7, es una regla que define el porcentaje de valores de una distribución normal que se encuentran a tres desviaciones estándar de la media.En concreto, la regla empírica establece lo siguiente:El 68% de los valores de una distribución normal se encuentran a una desviación estándar de la media.El 95% de los valores de una distribución normal se encuentran a dos desviaciones estándar de la media.El 99,7% de los valores de una distribución normal se encuentran a tres desviaciones estándar de la media. La distribución normal o distribución gaussiana es la distribución de probabilidad en variable continua, en la que la función densidad de probabilidad está descrita por una función exponencial de argumento cuadrático y negativo, que da lugar a una forma acampanada. El nombre de distribución normal viene del hecho que esta distribución es la que se aplica a mayor número de situaciones donde está involucrada alguna variable aleatoria continua en un grupo o población dada. Figura 1. Distribución normal $N(x; \mu,\sigma)$ y su densidad de probabilidad $f(s; \mu,\sigma)$. (Elaboración propia) Como ejemplos donde se aplica la distribución normal se tienen: la altura de los hombres o de las mujeres, variaciones en la medida de alguna magnitud física o en rasgos psicológicos o sociológicos medibles como el cociente intelectual o los hábitos de consumo de cierto producto. Por otra parte, se le llama distribución gaussiana o campana de Gauss, porque es a este genio matemático alemán a quién se le acredita su descubrimiento por el uso que le dio para la descripción del error estadístico de las mediciones astronómicas allá por el año 1800. Sin embargo, se afirma que esta distribución estadística fue publicada previamente por otro gran matemático de origen francés, como lo fue Abraham de Moivre, allá por el año 1733. [toc] Fórmula A la función distribución normal en la variable continua x, con parámetros μ y σ se le denota por: $N(x; \mu,\sigma)$ y explícitamente se escribe así: $N(x; \mu,\sigma) = \int_{-\infty}^{x} f(s; \mu,\sigma)\, ds$ donde $f(u; \mu,\sigma)$ es la función densidad de probabilidad: $f(s; \mu,\sigma) = (1/(\sigma\sqrt{2\pi})) \operatorname{Exp}(-s^2/(2\sigma^2))$ La constante que multiplica a la función exponencial en la función densidad de probabilidad se le llama constante de normalización, y se ha elegido de tal manera que: $N(+\infty, \mu,\sigma) = 1$ La expresión anterior asegura que la probabilidad de que la variable aleatoria x esté comprendida entre $-\infty$ y $+\infty$ sea 1, es decir el 100% de probabilidad. El parámetro μ es la media aritmética de la variable aleatoria continua x y σ la desviación típica o raíz cuadrada de la varianza de esa misma variable. En el caso que $\mu = 0$ y $\sigma = 1$ se tiene entonces la distribución normal estándar o distribución normal típica: $N(x; \mu = 0, \sigma = 1)$ Características de la distribucion normal 1- Si una variable estadística aleatoria sigue una distribución normal de densidad de probabilidad $f(s; \mu,\sigma)$, la mayor parte de los datos se agrupan alrededor de valor medio μ y están dispersos a su alrededor de forma tal que poco más de ⅔ de los datos están entre $\mu - \sigma$ y $\mu + \sigma$. Puede servirte: Prisma heptagonal2- La desviación típica σ siempre es positiva. 3- La forma de la función de densidad f se asemeja a la de una campana, por lo que a esta función muchas veces se le llama campana de Gauss o función gaussiana. 4- En una distribución gaussiana la media, la mediana y la moda coinciden. 5- Los puntos de inflexión de la función densidad de probabilidad se encuentran justamente en $\mu - \sigma$ y $\mu + \sigma$. 6- La función f es simétrica respecto a un eje que pase por su valor medio μ y tiene asintóticamente a cero para $x \longrightarrow +\infty$ y $x \longrightarrow -\infty$. 7- A mayor valor de σ mayor dispersión, ruido o distanciamiento de los datos alrededor del valor medio. Es decir a mayor σ la forma de campana es más abierta. En cambio σ pequeño indica que los dados se ciñen a la media y la forma de la campana es más cerrada o puntiaguda. 8- La función de distribución $N(x; \mu,\sigma)$ indica la probabilidad que la variable aleatoria sea menor o igual que x. Por ejemplo, en la figura 1 (más arriba) la probabilidad P de que la variable x sea menor o igual a 1.5 es de 84% y se corresponde con el área bajo la función densidad de probabilidad $f(x; \mu,\sigma)$ desde $-\infty$ hasta x. Intervalos de confianza 9- Si los datos siguen una distribución normal, entonces 68,26% de estos están entre $\mu - \sigma$ y $\mu + \sigma$. 10- El 95,44% de los datos que siguen una distribución normal se encuentran entre $\mu - 2\sigma$ y $\mu + 2\sigma$. 11- El 99,74% de los datos que siguen una distribución normal se encuentran entre $\mu - 3\sigma$ y $\mu + 3\sigma$. 12- Si una variable aleatoria x sigue una distribución $N(x; \mu,\sigma)$, entonces la variable $z = (x - \mu) / \sigma$ sigue la distribución normal estándar $N(z; 0,1)$. El cambio de la variable x a la z recibe el nombre de estandarización o tipificación y es de gran utilidad a al momento de aplicar las tablas de la distribución estándar a los datos que siguen una distribución normal no-estándar. Aplicaciones de la distribución normal Para aplicar la distribución normal es necesario pasar por el cálculo de la integral de la densidad de probabilidad, lo cual desde el punto de vista analítico no es fácil y no siempre se dispone de un programa informático que permita su cálculo numérico. Para este fin se usan las tablas de valores normalizados o tipificados, que no es más que la distribución normal en el caso $\mu = 0$ y $\sigma = 1$. Puede servirte: Tipos de ángulos, características y ejemplosTabla de distribución normal tipificada (parte 1/2) Tabla de distribución normal tipificada (parte 2/2) Debe notarse que estas tablas no incluyen los valores negativos. Sin embargo, usando las propiedades de simetría de la función densidad de probabilidad gaussiana pueden obtenerse los valores correspondientes. En el ejercicio resuelto mostrado más abajo se indica el uso de la tabla en estos casos. Ejemplo Suponga que tiene un conjunto de datos aleatorios x que siguen una distribución normal de media 10 y desviación típica 2. Se pide encontrar la probabilidad de que: a) La variable aleatoria x sea menor o igual a 8. b) Sea menor o igual a 10. c) Que la variable x esté por debajo de 12. d) La probabilidad que un valor x esté entre 8 y 12. Solución: a) Para responder a la primera pregunta simplemente hay que calcular: $N(x; \mu,\sigma)$ Con $x = 8$, $\mu = 10$ y $\sigma = 2$. Nos percatamos que se trata de una integral que no tiene una solución analítica en funciones elementales, sino la solución está expresada en función de la función error erf(x). Por otra parte, existe la posibilidad de resolver la integral en forma numérica, que es lo que hacen muchas calculadoras, las hojas de cálculo y programas informáticos como GeoGebra. La siguiente figura muestra la solución numérica correspondiente al primer caso: Figura 2. Densidad de probabilidad $f(x; \mu,\sigma)$. El área sombreada representa $P(x \leq 8)$. (Elaboración propia) y la respuesta es que la probabilidad que x esté por debajo de 8 es: $P(x \leq 8) = N(x=8; \mu=10,\sigma=2) = 0{,}1587$ b) En este caso se trata de encontrar la probabilidad de que la variable aleatoria x esté por debajo de la media que en este caso vale 10. La respuesta no requiere cálculo alguno, ya que sabemos que la mitad de los datos están por debajo de la media y la otra mitad por encima de la media. Por ello, la respuesta es: $P(x \leq 10) = N(x=10; \mu=10,\sigma=2) = 0{,}5$ c) Para responder a esta pregunta hay que calcular $N(x=12; \mu=10,\sigma=2)$, lo cual puede hacerse con una calculadora que tenga funciones estadísticas o mediante un software como es el caso de GeoGebra: Puede servirte: Eventos independientes: demostración, ejemplos, ejerciciosFigura 3. Densidad de probabilidad $f(x; \mu,\sigma)$. El área sombreada representa $P(x \leq 12)$. (Elaboración propia) La respuesta a la parte c puede verse en la figura 3 y es: $P(x \leq 12) = N(x=12; \mu=10,\sigma=2) = 0{,}8413$. d) Para encontrar la probabilidad de que la variable aleatoria x esté comprendida entre 8 y 12 podemos usar los resultados de las partes a y c de la siguiente manera: $P(8 \leq x \leq 12) = P(x \leq 12) - P(x \leq 8) = 0{,}8413 - 0{,}1587 = 0{,}6826 = 68{,}26\ \%$. Ejercicio resuelto El precio promedio de las acciones de una empresa es de \$25 con una desviación típica de \$4. Determine la probabilidad que: a) Una acción tenga un costo menor de \$20. b) Que tenga un costo mayor de \$30. c) El precio esté comprendido entre \$20 y \$30. Usar las tablas de distribución normal tipificada para encontrar las respuestas. Solución: Para poder hacer uso de las tablas, es necesario pasar a la variable z normalizada o tipificada: \$20 en la variable normalizada equivale a z = (\$20 – \$25) / \$4 = -5/4 = -1,25 y \$30 en la variable normalizada equivale a z = (\$30 – \$25) / \$4 = +5/4 = +1,25. a) \$20 equivale a -1,25 en la variable normalizada, pero la tabla no tiene valores negativos, por lo que ubicamos el valor +1,25 que arroja el valor de 0,8944. Si a este valor se le resta 0,5 el resultado será el área entre 0 y 1,25 que, por cierto, es idéntica (por simetría) al área entre -1.25 y 0. El resultado de la resta es 0,8944 – 0,5 = 0,3944 que es el área entre -1.25 y 0. Pero interesa el área desde $-\infty$ hasta -1,25 que será 0,5 – 0,3944 = 0,1056. Se concluye por tanto que la probabilidad que una acción esté por debajo de \$20 es 10,56%. b) \$30 en la variable tipificada z es 1,25. Para este valor en la tabla aparece el número 0,8944 que corresponde al área desde $-\infty$ hasta +1,25. El área entre +1.25 y $+\infty$ es (1 – 0,8944) = 0,1056. Es decir que la probabilidad que una acción cueste más de \$30 es 10,56%. c) La probabilidad de que una acción tenga un costo comprendido entre \$20 y \$30 se calculará así: 100% -10,56% – 10,56% = 78,88% Referencias Estadística y probabilidad. Distribución normal. Recuperado de: proyectodescartes.org Geogebra. Geogebra clásico, cálculo de probabilidad. Recuperado de geogebra.org MathWorks. Distribución de Gauss. Recuperado de: es.mathworks.com Mendenhall, W. 1981. Estadística para Administración y Economía. 3ra. edición. Grupo Editorial Iberoamérica. Stat Trek. Teach yourself Statistics. Poisson Distribution. Recuperado de: stattrek.com, Triola, M. 2012. Elementary Statistics. 11th. Ed. Pearson Education. Universidad de Vigo. Principales distribuciones continuas. Recuperado de: anapg.webs.uvigo.es Wikipedia. Distribución normal. Recuperado de: es.wikipedia.org La distribución normal es la más notable en la Estadística y el Cálculo de Probabilidad. A ella se ajustan infinidad de variables aleatorias continuas presentes en la naturaleza, las ciencias sociales, las finanzas o la biomedicina. Llamamos a esa característica: X. Queda definida por dos parámetros: la media y la desviación estándar (o desviación típica). Se denota por $N(\mu, \sigma)$. La función de densidad de probabilidad de la distribución normal Una función de densidad o de densidad de probabilidad $y = f(x)$ de una distribución de una variable aleatoria continua X es la que asigna a cada intervalo de la variable x_0, $x_0 + \Delta x$ la probabilidad de que X tome valores dentro de ese intervalo. La probabilidad no es el valor de la función en x, sino del área comprendida bajo la curva de la función en dicho intervalo). La función de densidad de probabilidad de la distribución normal es: Por su forma, a la gráfica de esta función de la distribución normal se le llama campana de Gauss. Características y propiedades Sus características y propiedades son: Tiene dos asíntotas en $+\infty$ y $-\infty$ en el eje de las abscisas. Es simétrica. En su eje de simetría coinciden la media, la mediana y la moda. El área entre la gráfica de la campana y el eje de las abscisas es la unidad. Representa la probabilidad (100 %) segura de se presenten todos los sucesos. En consecuencia, hay un 50 % de probabilidad de que un dato sea igual o mayor que la media y, en lógica, otro 50 % de que sea inferior a ella. En una distancia de $\mu \pm \sigma$ están los valores correspondientes a las abscisas de los dos puntos de inflexión de la campana de gauss. En este intervalo $\mu \pm \sigma$, bajo la curva hay un área de 0,6826. Representa el 68,26 % de posibilidades de encontrar un valor determinado en ese intervalo. En $\mu \pm 2\sigma$, se abarca un área de 0,9544, es decir, la probabilidad del 95,44 % de que allí esté un valor determinado. En $\mu \pm 3\sigma$, se cubre un área de 0,9974, es decir, la probabilidad del 99,74 % de que allí esté un valor determinado. En el intervalo $\mu \pm 1{,}96\sigma$, el área es de 0,95, es decir, la probabilidad del 95 % de que allí esté un valor determinado. La campana será más aplanada cuanto mayor sea la variabilidad de los datos, o lo que es lo mismo, cuanto mayor sea la desviación estándar. Y, al contrario, será más apuntada cuanto los datos estén más agrupados en torno a la μ (sea menor σ). El valor de la media hace que la campana se desplace a izquierda y derecha a lo largo del eje de las abscisas. No existe una única distribución normal, pues cada una está determinada por una media μ y una desviación estándar σ. Pero la 'forma' es común a todas ellas, simétricas y con un valor de exceso de curtosis igual a cero. Función de distribución acumulativa La función de distribución acumulativa F(x) o simplemente función de distribución es una función derivada que proporciona la probabilidad de que un valor x de la variable aleatoria X sea igual o menor que un valor determinado k. (La notación X se refiere a una variable aleatoria continua, mientras que x se refiere a cualquier valor real). F(x) representa el área comprendida bajo la gráfica para diferentes valores de la campana de Gauss. En cualquier función de distribución, sus valores extremos, en el límite son 0 y 1: Esta gráfica sería la de un histograma a partir de un polígono de frecuencias acumuladas, en el que los rectángulos tuviesen un ancho infinitesimal. Si se busca la probabilidad en un intervalo (k, l): La diferencia de los valores de la función de distribución entre los extremos del intervalo será la probabilidad buscada, como se ve en la imagen: Distribución normal estándar La distribución normal estándar o distribución normal tipificada es una distribución normal singular cuya denominación es N(0, 1). Su variable, Z es el producto de una transformación o cambio de variable de la variable X. Esta transformación se llama tipificación: La función de densidad de probabilidad de la distribución normal estándar o tipificada es: Con la tipificación, respecto a las gráficas, se produce un desplazamiento horizontal hacia el centro de coordenadas (0, 0) y un desplazamiento en la forma vertical (hacia arriba o hacia abajo): Es decir, de la distribución normal con variable X: Después de la tipificación se llega a la distribución normal estándar: Al tipificar una variable X y llegar a una distribución normal estandarizada, se consigue la comparación entre distribuciones diferentes, al tiempo que se facilita el cálculo de la probabilidad mediante el uso de una tabla normal estándar. La tabla que se ofrece a continuación proporciona directamente la probabilidad acumulada de que un suceso sea igual o menor que un valor positivo de Z: En los ejercicios se practicará con el manejo de la tabla. Ejercicios resueltos Ejercicio 1 El tiempo medio de duración de una batería de la flota de vehículos de una gran empresa es de cuatro años, con una desviación típica de medio año. ¿Qué probabilidad hay de que la batería de un vehículo escogido al azar haya durado igual o menos de cuatro años y medio? Solución: Este caso no requiere recurrir a estandarizar y uso de tablas, ya que se sabe que en una distribución normal, la probabilidad de encontrar un suceso comprendido entre la media y la desviación típica es la mitad del 68,26 %, o sea del 34,14 %. Por lo tanto, la probabilidad buscada será: Gráficamente: Habrá un 84,14 % de probabilidades de elegir al azar una batería que haya durado igual o menos de cuatro años y medio. Ejercicio 2 Un modelo de lámparas led destinado a la iluminación general de la vía pública de una gran ciudad tiene una vida útil media de 1750 horas y la desviación típica es de 800 horas. a) Hallar la probabilidad de que una bombilla cualquiera haya durado 2110 horas o menos. b) Qué probabilidad habrá de encontrar al azar un bombilla cuya duración haya sido superior a 1550 horas. c) En una partida de 2000 de estas bombillas, ¿Cuántas habrán durado entre 1310 y 2775 horas? Solución: a) Estandarizaremos el valor de 2110 para esta distribución: Buscamos este valor en la tabla. Vamos a la columna "0,05" de las centésimas. El valor de la probabilidad será: La intersección de fila y columna nos llevará a la casilla que corresponde a la probabilidad buscada. Su valor es 0,6736. Es decir, hay una probabilidad del 67,36 % de encontrar una bombilla que dure al menos 2110 horas. b) Estandarizaremos el valor de 1550 para esta distribución: En la tabla no hay valores negativos, no está -0,25. Pero la gráfica es simétrica respecto a la media, respecto al 0. Entonces, operaremos así: Igualdad que se comprueba en la imagen: Ya se puede buscar en la tabla el valor de la probabilidad acumulada correspondiente a 0,25 en la fila de las decenas 0,20 y la columna de las centenas 0,05: Que nos lleva a la casilla de la probabilidad buscada. Su valor es 0,5987. La probabilidad de encontrar al azar un bombilla cuya duración haya sido superior a 1550 horas es del 59,87 %. c) Se estandarizan los valores del intervalo: 1310 y 2775 horas: Como en esta tabla no hay valores de z negativos, para buscar la probabilidad acumulada a -0,55 se operará de forma similar al caso b). La gráfica es simétrica respecto a la media. Esta es la simetría. El área total comprendida bajo la campana es 1. Para ver la probabilidad correspondiente a un valor de z mayor que 0,55, se recurre a la tabla para ver la correspondiente a la acumulada hasta 0,55, dando por diferencia un resultado de 0,2912: El valor acumulado para 0,55 se ha obtenido de la tabla. Esa probabilidad acumulada es de 0,7088: Se aprecia en la imagen: La probabilidad correspondiente al valor negativo 0,55 es de 0,2912. Ahora se calcula la probabilidad hasta el valor del intervalo superior del problema, las 2775 horas, que estandarizado nos había dado una puntuación de z = 1,28. Lo buscaremos en la tabla: La tabla da una probabilidad para ese valor de 0,8997. La diferencia nos dará el valor de la probabilidad buscado para ese intervalo: La probabilidad de que una bombilla haya tenido una vida útil entre 1310 y 2775 horas es de 0,6085, o, lo que es lo mismo, del 60,85 %. El área sombreada bajo la curva normal tipificada se corresponde con esa probabilidad: Con esta probabilidad podremos calcular cuántas bombillas de esa partida de 2000 habrán durado entre 1310 y 2775 horas es de 2000 · 0,6085 = 1217 bombillas.