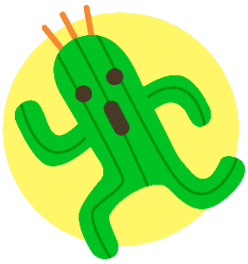
Estadística sin espinas

(3) Variables aleatorias



v0.1 23_11_09

Aprende sin espinas
con @carlos_cactus

A la inspiración del bucle_infinito,
al Cibergrupo y al tHash_A, por su amistad,
y sobre todo, a quienes dicen “pero quiero”
cuando sienten “no puedo”.

¡Un saludo sin espinas!
@carlos_cactus :D

Y si quieres saber más:
¡Encuéntrame en Telegram como @carlos_cactus o habla con Espinito, el bot Sin Espinas, en @GestionSinEspinasBot.

Únete a la comunidad de Telegram Sin Espinas y no te pierdas nada!

Deja de preocuparte por aprobar y ¡Aprende sin Espinas!

Variables aleatorias

VARIABLES ALEATORIAS

1.1. Definición de variable aleatoria

Una variable aleatoria es una función que asigna un valor numérico a cada elemento del espacio muestral.

Se define la variable aleatoria discreta X como una función que asigna valores numéricos a sucesos según:

$$X : \Omega \to \mathbb{R}$$

$$w_i \to X(w_i) = x_i$$

A cada suceso w_i del conjunto Ω se le asigna un valor x_i del conjunto $\mathbb{R}$.

Si se denota una variable aleatoria como X, la probabilidad con que X adopta un valor x se escribe como P(X = x).

Conocer la LEY o DISTRIBUCIÓN de una variable permite calcular cualquier probabilidad referente a su comportamiento.

Se modelan mediante funciones de distribución F(x) y además, según si son discretas o continuas:

- mediante funciones de MASA DE PROBABILIDAD p(x) en el caso de las variables DISCRETAS

- mediante funciones de DENSIDAD f(x) en el caso de las variables CONTINUAS.

Además, se caracterizan por 2 parámetros:

- La ESPERANZA matemática E(X) o VALOR ESPERADO.

- La VARIANZA Var(X).

Variables con la misma ley o distribución, comparten valores de E(X) y Var(X).

1.2. Tipos des de variables aleatorias

→ Variables aleatorias DISCRETAS

Pueden tomar valores finitos. Se destacan:

a) Distribución de Bernoulli
b) Distribución binomial
c) Distribución geométrica
d) Distribución de Poisson

Las variables aleatorias DISCRETAS se modelan mediante:

- Funciones de masa de probabilidad $p(x)$
- Funciones de distribución $F(x)$

→ Variables aleatorias CONTINUAS

Pueden tomar valores en un intervalo. Se destacan:

a) Distribución uniforme
b) Distribución exponencial
c) Distribución normal
d) T-Student
e) Chi cuadrado χ^2 (no se trata en este curso)

Las variables aleatorias CONTINUAS se modelan mediante:

- Función de densidad $f(x)$
- Funciones de distribución $F(x)$

1.3. Variables aleatorias DISCRETAS

1.3.1. DEFINICIÓN DE VARIABLE ALEATORIA DISCRETA

Pueden tomar valores finitos (o numerables, es decir, contable mediante números naturales) cuya probabilidad es siempre mayor que cero.

EJEMPLO

Se lanza al aire una moneda 3 veces. El conjunto de resultados posibles es:

Ω = { CCC , +CC , C+C , CC+ , ++C , +C+ , C++ , +++ }

Nótese la variación con repetición:
Se varían 2 elementos: "cara" = C y "cruz" = + , tomados de k en k.
Es decir, $N^k = 2^3$ posibilidades.

Se define la variable X como "número de caras".

Se observa:

X = x (número de caras)	Resultados	P(X=x)
0	+++	1/8
1	++C , +C+ , C++	3/8
2	+CC , C+C , CC+	3/8
3	CCC	1/8

Por ejemplo, se puede responder a cuál es la probabilidad de obtener más de 1 cara:

$$P(X > 1) = P\{(X = 2) \cup (X = 3)\} = P(X = 2) + P(X = 3) = \frac{3}{8} + \frac{1}{8} = \frac{4}{8}$$

EJEMPLO

Se lanzan 2 monedas.
Se define la variable aleatoria discreta X = "número de caras"

X Es una variable discreta ya que toma los valores en un conjunto numeral:
"1 cara", "2 caras"… no admite valores intermedios entre uno y otro.
Se denota "cara" como C y se denota "cruz" como X.
Se considera el conjunto de sucesos posibles:

$$\Omega = \{CC, CX, XC, XX\}$$

Se asigna a cada valor del conjunto Ω un valor de la variable x: $x : \Omega \to \mathbb{R}$ $\begin{cases} CC \to 2 \\ CX \to 1 \\ XC \to 1 \\ XX \to 0 \end{cases}$

1.3.2. FUNCIÓN DE MASA DE PROBABILIDAD P(X)

La función de masa de probabilidad p(x) es aquella que asigna a cada valor x del conjunto de origen de la variable X su probabilidad $P(X = x)$, es decir, un valor en el intervalo [0,1]:

$$p : x \rightarrow [0,1]$$
$$x_i \rightarrow p(X = x_i)$$

Esta función p(x) explica cómo se distribuye la probabilidad de los resultados de la variable X, o sea, la probabilidad con que X adopta cada valor x.

Para que p(x) esté bien definida, debe verificar 2 propiedades:

$$\sum_{i=1}^{n} p(x_i) = 1$$

- Es decir, que la suma de las probabilidades de todos los valores es 1.
- Asigna valores de probabilidad en el intervalo [0,1].

La función de probabilidad se suele definir a trozos.

EJEMPLO

Se lanzan 2 monedas.
Se define la variable X que modela el número de caras.
Se escribe la probabilidad de cada uno de los valores posibles, según los casos favorables de entre los posibles:

x_i	0	1	2
$p(x_i)$	1/4	2/4	1/4

Nótese que:

1) La suma de probabilidades es 1.
2) Todos los valores están en el intervalo [0,1].

Por tanto, p(x) está BIEN DEFINIDA.

Se puede escribir la función p(x) definida a trozos:

$$p(x) = \begin{cases} \frac{1}{4} & si\ x = 0 \\ \frac{1}{2} & si\ x = 1 \\ \frac{1}{4} & si\ x = 2 \end{cases}$$

1.3.3. FUNCIÓN DE DISTRIBUCIÓN DE UNA VARIABLE ALEATORIA DISCRETA

La función de distribución F(x) se basa en la acumulación de la función de probabilidad. Se denota como:

$$F(x_i) = P(X \leq x_i)$$

Es la función que asigna a cada valor x de la variable X la probabilidad con que adopta un valor igual o inferior a x (de ahí la ACUMULACIÓN).

Cumple 2 propiedades:

- Se define en el intervalo [0,1].
- Es SIEMPRE CRECIENTE.

EJEMPLO

Se desea calcular la función de distribución F(x) a partir de una tabla de probabilidades p dadas para una variable X que adopta valores x:

x_i	0	1	2	3
$p(x_i)$	0.1	0.2	0.4	0.3

→ Nótese de nuevo que para que la tabla describa una función de probabilidad debe cumplirse:

1. Los valores de probabilidad $p(x_i) \in [0,1]$.
2. Deben sumar 1: $\sum_{i=1}^{n} p(x_i) = 1$

En este caso, la función de probabilidad **está bien definida**:

- Los valores 0.1, 0.2, 0.4 y 0.3 están en el intervalo [0,1].
- Se cumple 0.1 + 0.2 + 0.4 + 0.3 = 1.

Dada la función de probabilidad, se pide calcular su función de distribución.

Para ello, se calcula la probabilidad acumulada para cada valor que adopta.

$$F(0) = P(X \leq 0) = P(X = 0) = 0.1$$

$$F(1) = P(X \leq 1) = P(X = 0) + P(X = 1) = 0.1 + 0.2 = 0.3$$

$$F(2) = P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = 0.1 + 0.2 + 0.4 = 0.7$$

$$F(3) = P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3);$$
$$; F(3) = 0.1 + 0.2 + 0.4 + 0.3 = 1$$

Además, la función de distribución se puede representar como tabla:

x_i	0	1	2	3
$F(x_i)$	0.1	0.3	0.7	1

EJEMPLO

Se lanza al aire una moneda 3 veces. Se observa:

X = x (número de caras)	Resultados	P(X=x)	$P(X \leq x)$
0	+++	1/8	1/8
1	++C , +C+ , C++	3/8	4/8
2	+CC , C+C , CC+	3/8	7/8
3	CCC	1/8	8/8

Se desea calcular el valor de la función de distribución F(x) para X = 2.9, es decir, F(2.9).

Se tiene:

$$F(2.9) = P(X \leq 2.9) = P(X \leq 2) = \frac{7}{8}$$

Es decir, en el intervalo de valores [0,2.9] se acumula 7/8 de la probabilidad.

1.3.4. INDEPENDENCIA DE VARIABLES ALEATORIAS DISCRETAS

Las variables aleatorias X e Y son independientes si se cumple:

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y)\ \forall x, y$$

Es decir, que para 2 valores x e y cualesquiera que adopten X e Y, la probabilidad con que adoptan x e y equivale al producto de probabilidades con que X adopta x por la probabilidad con que Y adopta y (igual que para los sucesos independientes).

Variables independientes verifican (igual que sucesos independientes):

$$P(X = x \mid Y = y) = P(X = x) \cap P(Y = y)$$

Siendo $P(X = x \mid Y = y)$ la probabilidad con que X adopta x si Y adopta y.

1.3.5. ESPERANZA E(X) Ó Μ DE UNA DISTRIBUCIÓN DISCRETA

La esperanza E(X), VALOR ESPERADO o MEDIA μ de una distribución discreta se define como:

$$E(X) = \mu = \sum_{i=1}^{n} x_i \cdot p(x_i) = \sum_{i=1}^{n} x_i \cdot P(X = x)$$

Es decir, es una media ponderada de los valores que adopta la variable, de acuerdo con la probabilidad con que lo hace.

EJEMPLO

A partir de la tabla de valores (nótese que se asimila la probabilidad a la frecuencia relativa):

x_i	0	1	2	3
$p(x_i)$	0.1	0.2	0.4	0.3

Se desea conocer el valor de la esperanza de la variable aleatoria X

$$E(X) = \mu = 0 \cdot 0.1 + 1 \cdot 0.2 + 2 \cdot 0.4 + 3 \cdot 0.3 = 0 + 0.2 + 0.8 + 0.9 = 1.9$$

EJEMPLO

Se tienen los datos:

Valores posibles X = x	38	60	27	12
P(X=x)	20/60	5/60	25/60	10/60

Se desea conocer el valor de la esperanza de la variable aleatoria X

Entonces:

$$E(X) = \mu = \sum_{i=1}^{n} x_i \cdot p(x_i) = 38 \cdot \frac{20}{60} + 60 \cdot \frac{5}{60} + 27 \cdot \frac{25}{60} + 12 \cdot \frac{10}{60} = 30.91\hat{6}$$

1.3.6. PROPIEDADES DE LA ESPERANZA E(X)

Tanto para distribuciones DISCRETAS como para distribuciones CONTINUAS, la esperanza E(X) cumple 3 propiedades:

1. La esperanza de una constante es esa misma constante:

   Una constante es una variable aleatoria que solo adopta el valor k con probabilidad 1. Entonces:

$$E(X) = \sum_{i=1}^{n} x_i \cdot P(X = x) = k \cdot P(X = k) = k \cdot 1 = k$$

2. La esperanza del producto de una variable aleatoria X por una constante k escala en un factor k la esperanza E(X):

$$E(kX) = kE(X)$$

3. La esperanza de la suma de variables aleatorias es la suma de sus esperanzas:

$$E(X + Y) = E(X) + E(Y)$$

   En el caso particular de una constante:

$$E(X + k) = E(X) + k$$

1.3.7. VARIANZA Var(X) σ^2 DE VARIABLES ALEATORIAS

La varianza Var(X) de una variable discreta X se define como:

$$Var(X) = E[X - E(X)]^2 = E(X^2) - E(X)^2$$

Es decir, la varianza es el valor esperado (promedio) de la desviación respecto de la media. Escrita en función de la frecuencia relativa:

$$Var(X) = \sigma^2 = \underbrace{\frac{1}{N}\sum_{i=1}^{n}(x_i - \mu)^2}_{\substack{de\ Estadísitica\ Descriptiva \\ DESACONSEJADA}} = \underbrace{\sum_{i=1}^{N} f_i(x_i - \bar{x})^2}_{En\ función\ de\ f_i}$$

Asimilando la frecuencia relativa f_i a la probabilidad, se tiene:

$$Var(X) = \sum_{i=1}^{n}\left(\mathrm{x_i} - \mathrm{E(X)}\right)^2 \cdot P(X = x_i)$$

Aunque la definición más PRÁCTICA es:

$$Var(X) = \sigma^2 = \underbrace{\left(\sum_{i=1}^{n} x_i^2 \cdot p(x_i)\right) - \mu^2}_{\substack{de\ Estadísitica\ Descriptiva \\ RECOMENDADA}} = \underbrace{E(X^2)}_{\substack{Promedio \\ de\ x_i^2}} - \underbrace{E(X)^2}_{\mu^2}$$

En definitiva:

$$Var(X) = \underbrace{E(X^2)}_{\substack{Promedio \\ de\ x_i^2}} - \underbrace{E(X)^2}_{\mu^2}$$

EJEMPLO

Se tiene:

x_i	0	1	2	3
$p(x_i)$	0.1	0.2	0.4	0.3

Se desea conocer el valor de la varianza de la variable aleatoria

La varianza Var(X) es:

$$Var(X) = \sigma^2 = \left(\sum_{i=1}^{n} x_i^2 \cdot p(x_i)\right) - \mu^2$$

En este caso:

$$Var(X) = (0^2 \cdot 0.1 + 1^2 \cdot 0.2 + 2^2 \cdot 0.4 + 3^2 \cdot 0.3) - 1.9^2 = 0.89$$

1.3.8. DESVIACIÓN TÍPICA σ DE UNA DISTRIBUCIÓN DISCRETA

Se define como:

$$\sigma = \sqrt{\sigma^2} = \sqrt{Var(X)}$$

Es una buena medida de la dispersión, ya que tiene las mismas unidades que la esperanza.

EJEMPLO

A partir del ejemplo anterior:

x_i	0	1	2	3
$p(x_i)$	0.1	0.2	0.4	0.3

Se tiene:

$$\sigma^2 = 0.89$$

Entonces, la desviación típica es:

$$\sigma = \sqrt{0.89} = 0.94$$

1.3.9. PROPIEDADES DE LA VARIANZA VAR(X)

Tanto para distribuciones DISCRETAS como para distribuciones CONTINUAS, la varianza Var(X) cumple 3 propiedades:

1. La varianza de una constante es 0 (puesto que una constante es una variable aleatoria que solo adopta un valor k con probabilidad 1):

$$Var(X) = E(X^2) - E(X)^2 = k^2 - k^2 = 0$$

2. La varianza del producto de una variable aleatoria X por una constante k equivale al producto de la varianza de X por el CUADRADO de k:

$$Var(kX) = k^2 Var(X)$$

De lo cual, para la desviación típica, se cumple:

$$\sigma(kX) = k\sigma(X)$$

Esta propiedad se demuestra:

$$Var(kX) = E[(kX)^2] - E(kX)^2 = k^2(E(X) - E(X)^2) = k^2 Var(X)$$

3. La varianza de la suma de variables aleatorias INDEPENDIENTES equivale a la suma de sus varianzas:

$$Var(X + Y) = Var(X) + Var(Y)$$

En el caso particular de una constante, nótese que Var(X) no varía:

$$Var(X + k) = Var(X)$$

1.3.10. DESIGUALDAD DE TCHEBICHEV

Provee una interpretación conjunta de la información que contienen la esperanza y la varianza.

Es VÁLIDA para variables aleatorias DISCRETAS y para CONTINUAS.

Se puede considerar que define:

1. La probabilidad MÁXIMA EXCLUIDA de intervalos de valores de longitud $2n\sqrt{Var(X)}$ centrados $E(X)$.

2. La probabilidad MÍNIMA ALOJADA en intervalos de valores de longitud $2n\sqrt{Var(X)}$ centrados $E(X)$.

Nótese que ambas propiedades son totalmente complementarias.

En definitiva, expresa con qué probabilidad un valor SE ESCAPA más allá de un cierto umbral respecto la esperanza, definido como un múltiplo de la desviación típica, o bien queda incluido DENTRO de ese umbral.

Para entenderlo:

- Se considera una variable X con valor esperado E(X) y varianza Var(X).

- Se define un intervalo de valores centrado en E(X), cuya longitud será de $2n\sqrt{Var(X)}$, es decir, 2n veces la desviación típica.

- Entonces, tanto a izquierda como a derecha de E(X) queda una longitud de n desviaciones típicas $n\sqrt{Var(X)}$.

La desigualdad de Tchebichev verifica que:

- DENTRO de ese intervalo se aloja una masa de probabilidad MÍNIMA equivalente a $1 - \frac{1}{n^2}$.

- FUERA de ese intervalo queda excluida una masa de probabilidad MÁXIMA de hasta $\frac{1}{n^2}$.

Es decir:

$$P\left(|X - E(X)| \geq n\sqrt{Var(X)}\right) \leq \frac{1}{n^2}$$

La probabilidad con que un valor adoptado por X se ESCAPE de un intervalo entre E(X) y una longitud de $n\sqrt{Var(X)}$ es, COMO MÁXIMO de HASTA $\frac{1}{n^2}$.

Y se afirma que un valor de X quedará DENTRO del intervalo con una probabilidad MÍNIMA de $1 - \frac{1}{n^2}$.

Nótese:

$$P\left(\underbrace{|X - E(X)|}_{\substack{Distancia\ entre\\ la\ variable\\ y\ su\ esperanza}} \underbrace{\geq}_{\substack{El\ valor\ de\ X\\ se\ ESCAPA\\ del\ intervalo}} \underbrace{n\sqrt{Var(X)}}_{\substack{Longitud\ del\ intervalo\\ a\ cada\ lado\ de\ E(X)}}\right) \leq \underbrace{\frac{1}{n^2}}_{\substack{Masa\ de\ probabilidad\\ EXCLUIDA\ del\ intervalo}}$$

- El VALOR ABSOLUTO $|X - E(X)|$ permite considerar valores a ambos lados de E(X).
- El signo ≥ denota que la probabilidad P declara refiere la probabilidad MÁXIMA con que el valor de X EXCEDE (se ESCAPA) del intervalo definido.
- El término $n\sqrt{Var(X)}$ define el LÍMITE DEL INTERVALO, es el punto de referencia: la distancia respecto E(X) con que se compara la distancia con que X se separa de su esperanza E(X).

O lo que es lo mismo, DENTRO del intervalo siguiente:

$$\left(E(X) - n\sqrt{Var(X)}, E(X) + n\sqrt{Var(X)}\right)$$

Se encuentra, como MÍNIMO, una masa de probabilidad por valor de:

$$1 - \frac{1}{n^2}$$

Por tanto, una vez definidos los extremos del intervalo, se puede afirmar:

$$P\left(E(X) - n\sqrt{Var(X)} < X < E(X) + n\sqrt{Var(X)}\right) \geq 1 - \frac{1}{n^2}$$

Es decir, que la probabilidad con que un valor SE ALOJA DENTRO en el intervalo definido es, COMO MÍNIMO (nótese ≥) de $1 - \frac{1}{n^2}$.

De modo que $\frac{1}{n^2}$ es la probabilidad MÁXIMA con que un valor se ESCAPA.
Y $1 - \frac{1}{n^2}$ es la probabilidad MÍNIMA con que un valor está DENTRO.

EJEMPLO

Se tiene una variable aleatoria X cuya esperanza es E(X) = 2 y cuya varianza es Var(X) = 4.

Se desea conocer con qué probabilidad adoptará un valor separado de la media por 3 veces la desviación típica.

Se conoce:
E(X) = 2
Var(X) = 4
n = 3

Se verifica la desigualdad:

$$P\left(|X - E(X)| \geq n\sqrt{Var(X)}\right) \leq \frac{1}{n^2}$$

En este caso:

$$P\left(|X - 2| \geq 3\sqrt{4}\right) \leq \frac{1}{3^2}$$

Es decir, que la probabilidad con que X adopta un valor que diste MÁS 3 veces la desviación típica respecto la esperanza, es decir la probabilidad MÁXIMA de que se ESCAPE del intervalo es de 1/9.

Por tanto, dentro del intervalo habrá, COMO MÍNIMO 8/9. Para comprobarlo, se escribe el intervalo:

$$\left(E(X) - n\sqrt{Var(X)}, E(X) + n\sqrt{Var(X)}\right)$$

En este caso:

$$\left(2 - 3 \cdot \sqrt{4}, 2 + 3 \cdot \sqrt{4}\right) = (-4,8)$$

Una vez definidos los límites del intervalo, se puede afirmar que la variable X adoptará un valor DENTRO del intervalo con una probabilidad MÍNIMA de 1 – 1/9 = 8/9:

$$P(-4 < X < 8) \geq \frac{8}{9}$$

1.4. Distribuciones DISCRETAS

1.4.3. DISTRIBUCIÓN DE BERNOULLI

1.4.3.1. Definición

Una variable aleatoria que X sigue una distribución de Bernoulli de parámetro p, se denota como:

$$X \sim B(p)$$

Esta distribución se define mediante el parámetro p que es la probabilidad de ÉXITO.

El concepto de ÉXITO o FRACASO es relativo al contexto: el éxito se puede definir como "que el cohete explote" o "que el experimento falle" o el fracaso como "que todo vaya bien" o que no se encuentre ningún defecto".

Permite calcular la probabilidad de un suceso que solo toma 2 resultados: ÉXITO o FRACASO, es decir, se define para variables dicotómicas.

A los 2 resultados se asigna un valor:

- 1 para el caso de éxito.
- 0 para su complementario, el de fracaso.

Su probabilidad respectiva se designa como:

- Probabilidad de ÉXITO p $\quad p = P(X = 1)$
- Probabilidad de FRACASO q $\quad q = P(X = 0) = 1 - p$

Donde $p \in (0,1)$.

Es decir, exige que el éxito 1 y el fracaso 0 sean COMPLEMENTARIOS.

1.4.3.2. Esperanza de una distribución de Bernoulli

Su esperanza es:

$$E(X) = 0 \cdot P(X = 0) + 1 \cdot P(X = 1) = 0 + p = p$$

1.4.3.3. Varianza de una distribución de Bernoulli

Su varianza es:

$$Var(X) = \sum_{i=1}^{n} \left(\mathrm{x_i} - \mathrm{E(X)}\right)^2 \cdot P(X = x_i)$$

Entonces:

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

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

$$Var(X) = p^2 - p^3 + (1 + p^2 - 2p)p = p^2 - p^3 + p + p^3 - 2p^2 = -p^2 + p$$

De lo cual:

$$Var(X) = p(1 - p)$$

1.4.3.4. Probabilidad de una distribución de Bernoulli

La probabilidad de obtener k ÉXITOS es:

$$p(X = k) = p^k \cdot (1 - p)^{1-k}$$

Si se realizan n pruebas INDEPENDIENTES de Bernoulli, los resultados adoptan una distribución BINOMIAL. Es decir, la distribución de Bernoulli es un caso particular de Binomial con una única prueba (n = 1).

EJEMPLO

Un jugador de baloncesto lanza un balón a canasta desde la línea de triple.

El resultado puede ser encestar = 1 o bien no encestar = 0.
Se realizan experimentos hasta determinar empíricamente que la probabilidad de encestar es 0.62.

La variable X = "número de canastas encestadas" se puede escribir como:

$$X \sim B(0.62)$$

Tomando como éxito con probabilidad p el resultado "encestar":

$$P(X = 1) = 0.62$$

Y como fracaso "no encestar":

$$P(X = 0) = 1 - p = 1 - 0.62$$

Cuya esperanza es:

$$E(X) = p = 0.62$$

Y con varianza:

$$Var(X) = 0.62 \cdot (1 - 0.62)$$

La probabilidad de observar 3 éxitos es:

$$p(X = k) = p^k \cdot (1 - p)^{1-k}$$

Es decir:

$$p(X = 3) = 0.62^3 \cdot (1 - 0.62)^{1-3}$$

EJEMPLO

Se tiene una variable $X \sim B\left(\frac{1}{3}\right)$ y se desea conocer la probabilidad de que adopte el valor x = 4, su esperanza y su varianza.

La probabilidad con que una variable binomial X con parámetro p = 1/3 adopta un valor x = 4

$$p(X = 4) = \left(\frac{1}{3}\right)^4 \cdot \left(1 - \frac{1}{3}\right)^{1-4} = \frac{1}{81} \cdot \left(\frac{2}{3}\right)^{-3} = \frac{1}{81} \cdot \frac{27}{8} = \frac{3}{72} = \frac{1}{24}$$

La esperanza de la variable es:

$$E(x) = p = \frac{1}{3}$$

La varianza de la variable es:

$$Var(x) = p \cdot q = \frac{1}{3} \cdot \left(1 - \frac{1}{3}\right) = \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}$$

1.4.4. DISTRIBUCIÓN BINOMIAL

1.4.4.1. Definición

La distribución binomial está integrada por n pruebas INDEPENDIENTES de experimentos de Bernoulli. Se denota como:

$$X \sim B(n, p)$$

Donde:

X es la variable estudiada que se modela mediante la distribución B, o sea, que sigue una distribución binomial.
n es el número de pruebas independientes que se observan.
p la probabilidad del suceso de ÉXITO.

Por tanto, se define mediante el parámetro p que es la probabilidad de ÉXITO y el tamaño muestral n: el número n de experimentos de Bernoulli que se realizan.

Es decir, la distribución de Bernoulli se puede entender como un caso particular de la binomial en que n = 1, solo se hace 1 prueba:

$$\underbrace{B(p)}_{Bernoulli} \equiv \underbrace{B(1, p)}_{\substack{Binomial \\ con\, n=1}}$$

Permite calcular la probabilidad de un suceso que solo toma 2 resultados: ÉXITO o FRACASO, es decir, se define para variables dicotómicas.

También se puede entender la binomial como la suma de variables de Bernoulli INDEPENDIENTES:

$$B(n, p) = X_1 + X_2 + \cdots + X_n = B_1(p) + B_2(p) + \cdots B_n(p)$$

Entonces, es fácil comprobar que la esperanza es n veces la esperanza de una distribución de Bernoulli:

$$E\big(B(n, p)\big) = E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) \ldots + E(X_n) = n \cdot p$$

Y que la varianza es n veces la varianza de una distribución de Bernoulli:

$$Var\big(B(n, p)\big) = Var(X_1 + X_2 + \cdots + X_n) = Var(X_1) + Var(X_1) + Var(X_n)$$

$$Var\big(B(n, p)\big) = n \cdot p(1 - p)$$

1.4.4.2. Probabilidad de una distribución binomial

La probabilidad de obtener k ÉXITOS para una variable que sigue una distribución binomial es:

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

Donde:

P(X=k)	es la probabilidad que la variable X adopte el valor k.
n	es el número de pruebas independientes.
k	es el valor que adopta X, de lo cual se calcula la probabilidad.
p	es la probabilidad del suceso de éxito.

Es decir, la probabilidad de una distribución binomial es una COMBINACIÓN de n experimentos de Bernoulli:

$$P(X = k) = \underbrace{\binom{n}{k}}_{Combinación} \cdot \underbrace{p^k \cdot (1-p)^{n-k}}_{\substack{Probabilidad \\ de\ k\ éxitos\ en \\ experimento\ de \\ Bernoulli}}$$

1.4.4.3. APÉNDICE: Número combinatorio

Cabe recordar la definición del número combinatorio:

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

EJEMPLO

$$\binom{6}{4} = \frac{6!}{4! \cdot (6-4)!} = \frac{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{4 \cdot 3 \cdot 2 \cdot 1 \cdot (2 \cdot 1)} = \frac{6 \cdot 5}{2} = 15$$

En calculadora:

Pulsar: 6 -> SHIFT -> nCr (función secundaria en tecla DIVISIÓN) -> 4
Aparece en pantalla: 6 C 4
Nótense algunas propiedades:

$$0! = 1$$

$$\binom{n}{0} = \frac{n!}{0!\,n!} = \frac{1}{1} = 1$$

$$\binom{n}{n} = \frac{n!}{n!\,0!} = \frac{1}{1} = 1$$

$$\binom{n}{1} = \frac{n!}{1!\,(n-1)!} = \frac{n \cdot (n-1)!}{1! \cdot (n-1)!} = n$$

$$\binom{n+1}{n} = \frac{(n+1)!}{n!\,(n+1-n)!} = \frac{(n+1) \cdot n!}{n! \cdot 1!} = n+1$$

1.4.4.4. Esperanza de una distribución binomial

La esperanza de una distribución binomial es n veces la esperanza de una distribución de Bernoulli:

$$E(X) = \mu = n \cdot p$$

Donde:
n número de pruebas independientes.
p probabilidad del suceso de éxito.

1.4.4.5. Varianza de una distribución binomial

La varianza de una distribución binomial es n veces la varianza de una distribución de Bernoulli:

$$Var(X) = \sigma^2 = n \cdot p \cdot (1 - p)$$

Donde:
n número de pruebas independientes.
p probabilidad del suceso de éxito.

También se puede escribir en función de la probabilidad q de fracaso:

$$Var(X) = n \cdot p \cdot \underbrace{(1 - p)}_{q} = n \cdot p \cdot q$$

Donde
q probabilidad del suceso contrario a p, que cumple: q = 1 – p

1.4.4.6. Desviación típica de una distribución binomial:

Se cumple:

$$\sqrt{Var(X)} = \sigma = \sqrt{n \cdot p \cdot (1 - \mathrm{p})} = \sqrt{n \cdot p \cdot q}$$

1.4.4.7. Aproximación de variables binomiales a Normal

ES MUY RELEVANTE

Ver en apartado Teorema del Límite Central.

1.4.4.8. Aproximación de variables binomiales a Poisson

Ver en apartado Distribución de Poisson

Dada una binomial $X \sim B(n, p)$ que satisfaga:

1. $p < 0.1$
2. $np \leq 5$

Mediante el Teorema Central del Límite, se puede aproximar a una distribución de Poisson $Y \sim P(n \cdot p)$ de modo que el cálculo de probabilidad de la binomial Poisson se ajuste al cálculo de probabilidad de la Poisson Y.

$$P(X = k) = \underbrace{\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}}_{Binomial\ X} \approx \underbrace{\frac{(np)^k}{k!} e^{-np}}_{\substack{Poisson\ Y \\ P(Y=k)}}$$

EJEMPLO

Considérese un experimento cuya probabilidad de éxito es de 1/3 del cual se realizan 6 repeticiones. ¿Qué probabilidad hay de tener éxito en 4 de ellas?

Se tiene una variable $X \sim B\left(6, \frac{1}{3}\right)$ y se desea calcular la probabilidad asociada a x = 4.

$$p(X = 4) = \binom{6}{4} \left(\frac{1}{3}\right)^4 \cdot \left(1 - \frac{1}{3}\right)^{6-4} = 15 \cdot \frac{1}{81} \cdot \left(\frac{2}{3}\right)^2 = \frac{15}{81} \cdot \frac{4}{9} = \frac{60}{729}$$

La esperanza de la variable es:

$$E(x) = n \cdot p = 6 \cdot \frac{1}{3} = 2$$

La varianza de la variable es:

$$Var(x) = n \cdot p \cdot q = 6 \cdot \frac{1}{3} \cdot \left(1 - \frac{1}{3}\right) = 2 \cdot \frac{2}{3} = \frac{4}{3}$$

EJEMPLO

Se lanza una moneda al aire 6 veces. Se define X como el número de caras obtenidas. Se sabe que la probabilidad de obtener cara EQUIVALE a la de obtener cruz.

→ Se desea calcular la probabilidad de obtener 0 caras.

Se tiene un experimento de Bernoulli (solo puede salir cara o cruz) realizado en n = 6 pruebas INDEPENDIENTES. Por tanto, se puede modelar la variable X como una distribución binomial:

$$X = número\ de\ caras\ obtenidas$$

$$X \sim B(n, p) \rightarrow X \sim \left(6, \frac{1}{2}\right)$$

Con $n = 6\ lanzamientos$

Entonces:

$$p = \frac{1}{2} \rightarrow q = 1 - p = 1 - \frac{1}{2} = \frac{1}{2}$$

Se aplica:

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

Tomando: k = 0
n = 6
p = 0.5

Se obtiene:

$$P(X = 0) = \binom{6}{0} \cdot \left(\frac{1}{2}\right)^0 \cdot \left(1 - \frac{1}{2}\right)^{6-0} = 1 \cdot 1 \cdot \left(\frac{1}{2}\right)^6 = \frac{1}{64}$$

La probabilidad de obtener exactamente 0 caras en 6 lanzamientos es de 1/64.

→ Se desea calcular la probabilidad de obtener 6 caras:

$$P(X = 6) = \binom{6}{6} \cdot \left(\frac{1}{2}\right)^6 \cdot \left(1 - \frac{1}{2}\right)^{6-6} = 1 \cdot 1 \cdot \left(\frac{1}{2}\right)^6 = \frac{1}{64}$$

La probabilidad de obtener exactamente 6 caras en 6 lanzamientos es 1/64.

→ Se desea calcular la probabilidad de obtener 2 caras:

$$P(X = 2) = \binom{6}{2} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(1 - \frac{1}{2}\right)^{6-2} = 1 \cdot 1 \cdot \left(\frac{1}{2}\right)^4 = \frac{1}{16}$$

La probabilidad de obtener exactamente 2 caras en 6 lanzamientos es 1/16.

→ Se desea calcular la esperanza:

Se cumple que E(X) es:

$$\mathrm{E(X)} = \mu = n \cdot p = 6 \cdot \frac{1}{2} = 3$$

→ Se desea calcular la varianza:

Se cumple que Var(X) es:

$$\mathrm{Var(X)} = \sigma^2 = n \cdot p \cdot q = 6 \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{2}\right) = 3 \cdot \frac{1}{2} = \frac{3}{2}$$

→ Se desea calcular la desviación típica:

La desviación típica es la raíz cuadrada de la varianza:

$$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{6 \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{2}\right)} = \sqrt{\frac{3}{2}} = \frac{\sqrt{6}}{2}$$

EJEMPLO

Se dispone de una urna con bolas de dos colores: unas son blancas y otras son negras. Hay 6 bolas blancas y 8 negras. El experimento es extraer 5 bolas CON REEMPLAZAMIENTO (cada vez que se extrae una bola, se vuelve a introducir en la urna).

→ Se dese saber qué es más probable, ¿sacar 2 bolas blancas o 3 negras?

Se define:
X = número de bolas BLANCAS extraídas

Se tiene:
n = 5
$p = \frac{6}{14}$

La probabilidad de BLANCA es el ÉXITO, acorde a la definición de X:

Entonces:

$$X \sim B(n,p) \rightarrow X \sim B\left(5, \frac{6}{14}\right)$$

Se calcula la probabilidad de extraer 2 BLANCAS:

$$P(X=2) = \binom{5}{2} \cdot \left(\frac{6}{14}\right)^2 \cdot \left(\frac{8}{14}\right)^{5-2} = 0.3427$$

Para la extracción de bolas NEGRAS se define una nueva variable:

Y = número de bolas NEGRAS extraídas
n = 5
p = $\frac{8}{14}$ Se toma la probabilidad de NEGRA como ÉXITO, acorde con Y.
$q = 1 - \frac{8}{14} = \frac{6}{14}$

Entonces:

$$Y \sim B(n, p) \rightarrow Y \sim B\left(5, \frac{8}{14}\right)$$

Se calcula la probabilidad de extraer 3 NEGRAS:

$$P(Y = 3) = \binom{5}{3} \cdot \left(\frac{8}{14}\right)^3 \cdot \left(\frac{6}{14}\right)^{5-3} = 0.3427$$

Nótese que las dos expresiones, para BLANCAS y para NEGRAS, son equivalentes:

$$P(X = 2) = \binom{5}{2} \cdot \left(\frac{6}{14}\right)^2 \cdot \left(\frac{8}{14}\right)^3 = \binom{5}{3} \cdot \left(\frac{8}{14}\right)^3 \cdot \left(\frac{6}{14}\right)^2 = P(Y = 3)$$

Esto se debe a que $\binom{5}{2} = \binom{5}{3}$, ya que:

$$\binom{c}{a} = \binom{c}{b} \leftrightarrow a + b = c \rightarrow \frac{c!}{a!\,(c-a)!} = \frac{c!}{b!\,(c-b)!} \leftrightarrow \begin{cases} (c-a)! = b! \\ (c-b)! = a! \end{cases}$$

Es decir: $\binom{n}{m} = \binom{n}{n-m}$

1.4.4.9. Suma de binomiales con misma p

Una propiedad útil de la distribución binomial es que si varias variables que siguen distribución binomial comparten la misma probabilidad de éxito:

$$X \sim B(n, p), Y \sim B(m, p)$$

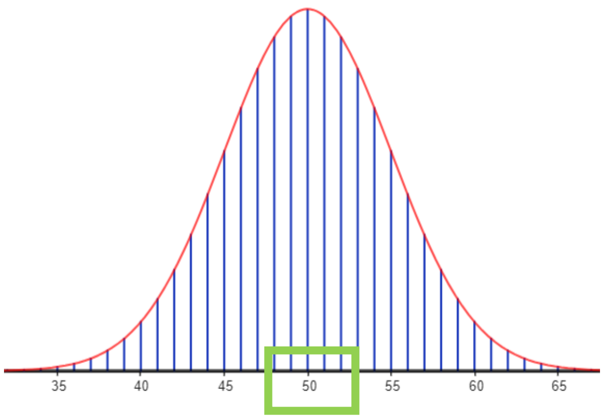
Su suma también presenta el mismo parámetro p:

$$X + Y \sim B(n + m, p)$$

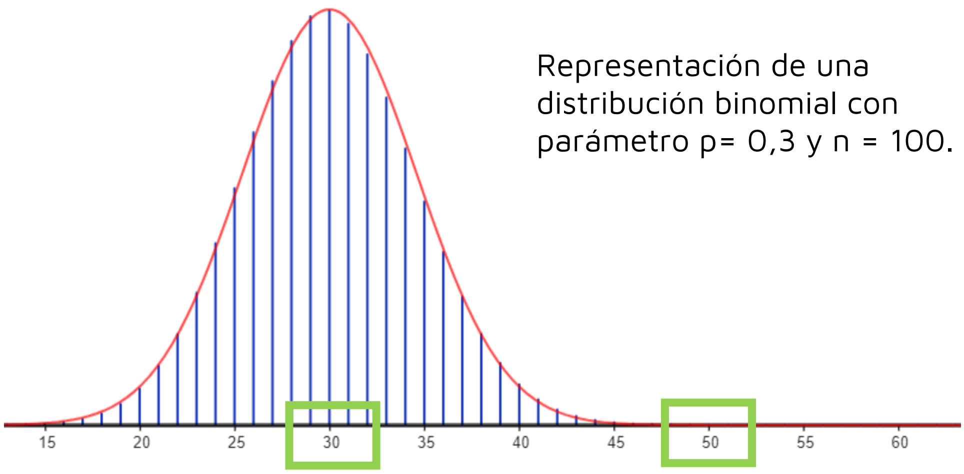
Esto radica en que la distribución binomial es una combinación de n experimentos de Bernoulli independientes.

1.4.4.10. Simetría de la distribución binomial

La distribución binomial es simétrica cuando p = 0,5:



Representación de una distribución binomial con parámetro p= 0,5 y n = 100.



Representación de una distribución binomial con parámetro p= 0,3 y n = 100.

1.4.5. DISTRIBUCIÓN UNIFORME DISCRETA

1.4.5.1. Definición

Se denota por:

$$X \sim U(N)$$

Se caracteriza por el número N de resultados distintos que puede adoptar. Verifica que todos los n resultados que puede tomar son equiprobables.

1.4.5.2. Función de masa de probabilidad

Dado que todos los resultados son equiprobables, se tiene:

$$p(X = k) = \frac{1}{N}$$

1.4.5.3. Esperanza de una distribución uniforme

Su esperanza E(x) es:

$$E(X) = \frac{N+1}{2}$$

1.4.5.4. Varianza de una distribución uniforme

Su varianza Var(x) es:

$$Var(X) = \frac{N^2 - 1}{12}$$

EJEMPLO

El lanzamiento de un dado de 6 caras es un buen ejemplo de distribución uniforme discreta: los resultados son discretos y equiprobables.

La probabilidad de obtener un 5 al lanzar un dado es:

$$p(X = 5) = \frac{1}{6}$$

La esperanza de la variable es:

$$E(X) = \frac{N+1}{2} = \frac{6+1}{2} = 3.5$$

La varianza de la variable es:

$$Var(X) = \frac{N^2 - 1}{12} = \frac{36 - 1}{12} = \frac{35}{12}$$

1.4.6. DISTRIBUCIÓN GEOMÉTRICA

1.4.6.1. Definición

Se aplica a la gestión de calidad: suele definirse el ÉXITO como encontrar un objeto DEFECTUOSO en una serie de inspecciones.
Se denota como:

$$X \sim Geom(p)$$

Se define mediante el parámetro p de la probabilidad del suceso de éxito.
La variable X mide:

- El **número de fracasos ANTES de obtener el PRIMER éxito**, es decir, el número de REPETICIONES hasta hallar el éxito.

- El número de intentos que hay que hacer HASTA el primer éxito.

1.4.6.2. Cálculo de probabilidad

La función de masa de probabilidad de una variable que sigue distribución geométrica es:

$$p(X = k) = \underbrace{q^{k-1}}_{\substack{k-1 \\ fracasos}} \cdot \underbrace{p}_{\substack{éxito \\ único}} = (1-p)^{k-1} \cdot p$$

Siendo:
q La probabilidad de fracaso q = 1 – p.
k El número MÁXIMO de fracasos ANTES del primer éxito.

Entonces, se observarán:

- k – 1 FRACASOS ANTES del primer éxito.
- El PRIMER éxito se obtiene en el intento k-ésimo.

1.4.6.3. Esperanza de una distribución geométrica

La esperanza de una variable que sigue una distribución geométrica es:

$$E(x) = \frac{1}{p}$$

1.4.6.4. Varianza de una distribución geométrica

La varianza de una variable que sigue una distribución geométrica es:

$$Var(x) = \frac{q}{p^2} = \frac{1-p}{p^2}$$

1.4.6.5. Función de distribución de una distribución geométrica

Se cumple:

$$P(X < k) = \mathrm{P}(\mathrm{X} \leq \mathrm{k} - 1) = 1 - (1 - \mathrm{p})^{\mathrm{k}-1} = 1 - \ q^{k-1}$$

Donde p es la probabilidad de ÉXITO, es decir, ENCONTRAR EL DEFECTO y q es la probabilidad de FRACASO, es decir, seguir sin encontrarlo.

EJEMPLO

Se tiene una variable X que representa el número de artículos seleccionados ANTES de encontrar uno defectuoso de acuerdo con la distribución:

$$X \sim Geom(0.01)$$

Se desea calcular la probabilidad de que el 5° artículo sea defectuoso, su esperanza y su varianza.

Se define:

ÉXITO	encontrar un artículo DEFECTUOSO (asociado a p).
FRACASO	encontrar uno EN BUEN ESTADO (asociado a q = 1 – p).

En este caso, si el 5° es DEFECTUOSO, se FRACASA 4 veces.

Se impone k = 5:

- 4 fracasos 4 primeros intentos
- 1 acierto 5° intento

Se desea conocer:

$$p(X = 5)$$

Se cumple:

$$p(X = k) = q^{k-1} \cdot p$$

Entonces:

$$p(X = 5) = (1 - 0.01)^{5-1} \cdot 0.01 = 9.605 \cdot 10^{-3}$$

La esperanza de la variable es:

$$E(x) = \frac{1}{p} = \frac{1}{0.01} = 100$$

Lo cual significa que la PRIMERA PIEZA DEFECTUOSA se espera en la posición 100.

La varianza de la variable es:

$$Var(x) = \frac{q}{p^2} = \frac{1 - 0.01}{0.01^2} = \frac{0.99}{0.01} = 9900$$

EJEMPLO

En un proceso de fabricación se produce una pieza NO defectuosa con una probabilidad de 0.95. Se desea conocer qué probabilidad hay de obtener 1 pieza defectuosa ANTES de llegar a examinar 10 piezas.

Se define la variable X que modela la POSICIÓN DE LA PRIMERA pieza DEFECTUOSA. Se cumple entonces:

$$X \sim Geom(0.05)$$

Siendo:

p = 0.05 Probabilidad de ÉXITO: es decir, que SEA DEFECTUOSA.
q = 0.95 Probabilidad de FRACASO: es decir, NO defectuosa.

Se impone k = 10 (10 intentos).
Se desea conocer:

$$P(X < 10)$$

Es decir, la probabilidad con que la primera pieza defectuosa se encuentre ANTES del décimo objeto analizado.

Se aplica la definición de la función de distribución:

$$P(X < k) = P(X \leq k-1) = F(k) = 1 - (1-p)^{k-1}$$

En este caso:

$$P(X < 10) = P(X \leq 9) = F(9) = 1 - (1 - 0.05)^{10-1} = 1 - 0.95^9$$

1.4.7. DISTRIBUCIÓN DE POISSON

1.4.7.1. Definición

Se denota como:

$$X \sim P(\lambda)$$

Se define mediante el parámetro λ que es la esperanza.

La distribución de Poisson se suele aplicar a modelar la frecuencia con que se da un suceso (número de veces en cada unidad de tiempo).

1.4.7.2. Función de masa de probabilidad

La distribución de Poisson se ajusta a la función de masa de probabilidad:

$$p(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

1.4.7.3. Esperanza de la distribución de Poisson

La esperanza de la variable es:

$$E(x) = \lambda$$

1.4.7.4. Varianza de la distribución de Poisson

La varianza de la variable es:

$$Var(x) = \lambda$$

1.4.7.5. Aproximación de variables binomiales a variables de Poisson

Dada una binomial $X \sim B(n, p)$ que satisfaga:

1. $p < 0.1$
2. $np \leq 5$

Si se dan estas 2 condiciones, mediante el Teorema Central del Límite, se puede aproximar la binomial $X \sim B(n, p)$ a una distribución de Poisson $Y \sim P(n \cdot p)$ de modo que el cálculo de probabilidad de la binomial se ajuste al cálculo de probabilidad de la Poisson Y.

Es decir:

$$P(X = k) = \underbrace{\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}}_{Binomial\ X} \approx \underbrace{\frac{(np)^k}{k!} e^{-np}}_{\substack{Poisson\ Y \\ P(Y=k)}}$$

EJEMPLO

Se lanzan 2 dados a la vez. Se hacen 60 tiradas. Se llama X = "número de veces que sale doble 6 en una tirada".

Se puede modelar X como una binomial:

$$X \sim B(n, p) \rightarrow X \sim B\left(60, \frac{1}{6^2}\right)$$

Se desea conocer la probabilidad de obtener doble 6 en 15 ocasiones. Para una binomial, se tendría que operar según:

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

En este caso:

$$P(X = 15) = \binom{60}{15} \cdot \left(\frac{1}{6^2}\right)^{15} \cdot \left(1 - \frac{1}{6^2}\right)^{60-15}$$

Lo cual es MUY LENTO.
Pero se observa que se cumple:

1. $p = \frac{1}{6^2} \ll 0.1$
2. $np = 60 \cdot \frac{1}{6^2} \ll 5$

Entonces, se puede aproximar X a una variable que sigue una distribución de Poisson:

$$X \sim B\left(60, \frac{1}{6^2}\right) \approx Y \sim P(np)$$

En este caso:

$$Y \sim P\left(60 \cdot \frac{1}{6^2}\right)$$

Lo cual permite calcular la probabilidad deseada de forma ágil:

$$P(X = k) = \underbrace{\binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}}_{Binomial\ X} \approx \underbrace{\frac{(np)^k}{k!} e^{-np}}_{\substack{Poisson\ Y \\ P(Y=k)}}$$

En este caso:

$$\underbrace{P\,(X = 15) = \binom{60}{15} \cdot \left(\frac{1}{6^2}\right)^{15} \cdot \left(1 - \frac{1}{6^2}\right)^{60-15}}_{Binomial\ X} \approx \underbrace{P(Y = 15) = \frac{\left(60 \cdot \frac{1}{6^2}\right)^{15}}{15!} e^{-60 \cdot \frac{1}{6^2}}}_{Poisson\ Y}$$

EJEMPLO

Sea X el número de camiones que llega a un almacén diariamente.
Se sabe que de media llegan 3 camiones al almacén diariamente.
Se desea calcular la probabilidad de que un día lleguen más 5 camiones.

Se tiene $X \sim P(3)$.
En este caso, la probabilidad de que lleguen MÁS de k = 5 camiones será:

$$P(X > 5) = P(X = 6) + P(X = 7) + P(X = 8) + \cdots + P(X = n)$$

O lo que es lo mismo, el caso complementario que permite el cálculo finito:

$$p(X > 5) = 1 - p(X \leq 5)$$

Es decir:

$$P(X > 5) = 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)]$$

En este caso con $\lambda = 3$:

$$P(X > 5) = 1 - \left(e^{-3} \cdot \frac{3^0}{0!} + e^{-3} \cdot \frac{3^1}{1!} + e^{-3} \cdot \frac{3^2}{2!} + e^{-3} \cdot \frac{3^3}{3!} + e^{-3} \cdot \frac{3^4}{4!} + e^{-3} \cdot \frac{3^5}{5!}\right)$$

Es decir:

$$P(X > 5) = 1 - e^{-3}\left(1 + \frac{3}{1} + \frac{9}{2} + \frac{27}{6} + \frac{81}{24} + \frac{243}{120}\right) = 0.0839179$$

La esperanza de la variable es:

$$E(x) = 3$$

La varianza de la variable es:

$$Var(x) = 3$$

EJEMPLO

Se sabe que el número de personas que pasa por la caja 1 de un supermercado en 1 hora sigue una distribución de Poisson con parámetro 15.

Se ha definido una variable X que cuenta el número de personas que pasa por esa caja.
Se desea conocer:

→ La probabilidad con que pasan EXACTAMENTE 12 personas por esa caja en 1 hora.

Se tiene:

$$X \sim P(15)$$

Se cumple:

$$P(X = \mathrm{k}) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

En este caso:

$$P(X = 12) = e^{-15} \cdot \frac{15^{12}}{12!}$$

→ La probabilidad de que pase ALGUNA persona en 1 hora.
Es decir:

$$P(X > 0)$$

Lo cual implicaría sumar:

$$P(X > 0) = P(X = 1) + P(X = 2) + P(X = 3) \ldots$$

Para superar el carácter NO FINITO de esa suma, se apela a la COMPLEMENTARIEDAD:

$$P(X > 0) = 1 - \mathrm{P(X = 0)} = 1 - e^{-15} \cdot \frac{15^0}{0!} = 1 - e^{-15}$$

→ El número esperado de personas en 6 horas.

Si en 1 hora se esperan 15, en 6 horas se esperarán 6·15 = 80 personas. Esto resulta del escalado de la muestra, que afecta en el mismo factor a la media, es decir, la esperanza, que es el parámetro λ de la variable.

1.6. Variables aleatorias continuas

1.6.1. DEFINICIÓN DE VARIABLE ALEATORIA CONTINUA

Se denominan variables aleatorias continuas o ABSOLUTAMENTE continuas a aquellas variables aleatorias que pueden tomar cualquier valor en un intervalo (como el peso o la altura de una persona, o la velocidad con que se mueve un animal).

Se define la variable aleatoria continua x como una función DE DENSIDAD (y no de masa de probabilidad, como en el caso de las discretas), habitualmente denotada como f(x).

1.6.2. DISTRIBUCIONES ALEATORIAS CONTINUAS

Se destacan:

1. Distribución uniforme
2. Distribución exponencial
3. Distribución normal
4. T-Student
5. Chi cuadrado χ^2 (no se trata en este curso)

Se modelan mediante:

- Función de densidad f(x)
- Funciones de distribución F(x)

Las funciones de distribución F(x) son las primitivas de las funciones de densidad:

$$F(x) = \int f(x)\ dx$$

1.6.3. FUNCIÓN DE DENSIDAD

Las funciones de densidad f(x) permiten definir variables aleatorias continuas.
La primitiva de la función de densidad f(x) es la función de distribución F(x).
La función de densidad de probabilidad debe cumplir 3 propiedades:

1. $f(x_i) \geq 0$ (nótese para variables aleatorias discretas: $p(x_i) \geq 0$)

   Es decir, solo considera probabilidades en [0,1] y NUNCA negativas.

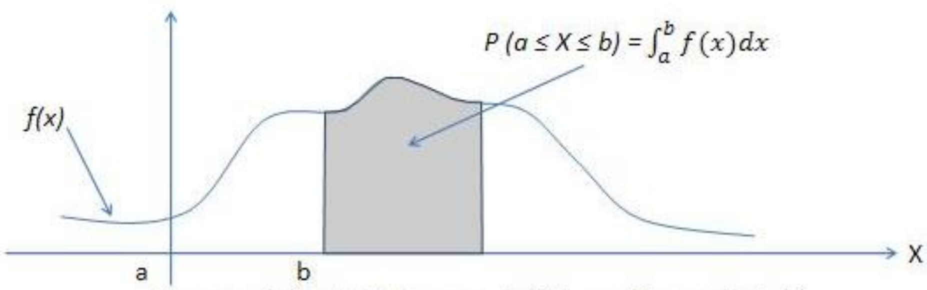
2. La probabilidad con que X adopta un valor en el intervalo [a,b] coincide con el área cerrada debajo de la curva que describe la función de densidad f(x), es decir, con el valor de su integral, denominada función de distribución F(x), evaluada en ese intervalo [a,b]:

$$\Big[\underbrace{F(x)}_{\substack{Función\ de \\ DISTIRBUCIÓN}}\Big]_a^b = p(a \leq X \leq b) = \int_a^b \underbrace{f(x)}_{\substack{Función\ de \\ DENSIDAD}} dx$$

Si se considera la probabilidad como el área cerrada por la función de distribución f(x), se tiene:

a) $f(x) \geq 0$, para toda x IR b) $\int_{-\infty}^{+\infty} f(x)dx = 1$

Así, para cualesquier reales a y b, tales que a ≤ b , tenemos $P(a \leq X \leq b) = \int_a^b f(x)\,dx$.



Muestra el área bajo la curva de *f(x)*, en el intervalo (a,b)

Es MUY RELEVANTE notar que, en una función continua, la PROBABILIDAD de un valor PUNTUAL, discreto, ES NULA, ya que la probabilidad acumulada en un punto es 0, es decir:

$$p(X = c) = 0$$

Por tanto, la probabilidad solo adopta valores no nulos en un INTERVALO.

3. $\int_{-\infty}^{\infty} f(x)dx = 1$ (para variables aleatorias discretas: $\sum p(x_i) = 1$)

El área total encerrada bajo la curva de la función de densidad es 1.

EJEMPLO

Dada la función de densidad f(x):

$$f(x) = \begin{cases} 2x & 0 < x < 1 \\ 0 & (-\infty, 0] \cup [1, +\infty) \end{cases}$$

Se desea calcular la función de distribución F(x).

Basta con integrar en cada una de las partes del dominio hasta el valor x:

$$F(x) = p(X \leq x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{0} f(x)dx + \int_{0}^{x} f(x)dx = 0 + [x^2]_0^x = x^2$$

Por tanto:

$$F(x) = \begin{cases} x^2 & 0 < x < 1 \\ 0 & (-\infty, 0] \cup [1, +\infty) \end{cases}$$

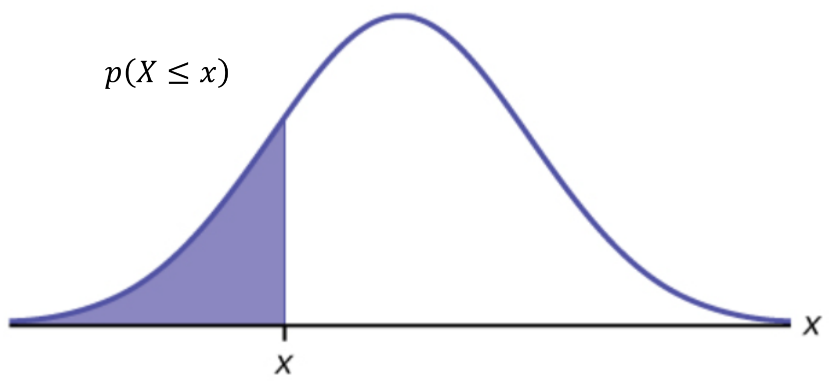
1.6.4. CÁLCULO DE PROBABILIDAD MEDIANTE FUNCIÓN DE DISTRIBUCIÓN F(X)

1. En primer lugar, nótese que como $p(X = c) = 0$ (la probabilidad acumulada en un punto es nula) se cumple:

$$p(X \leq c) = \mathrm{p(X < C)}$$

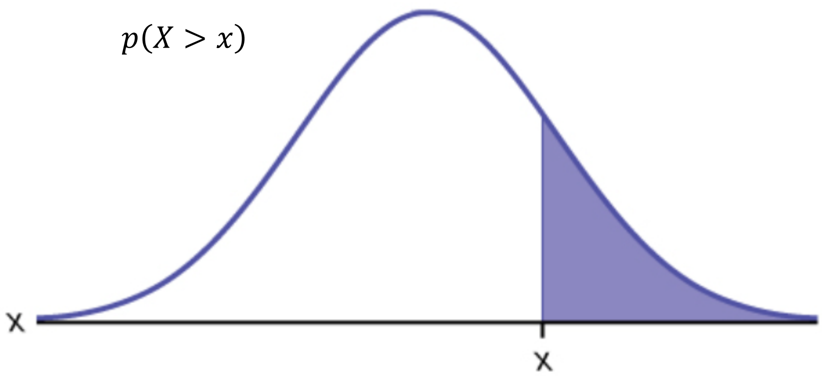
2. Para el cálculo de $p(X \leq a)$ se toma $F(a)$:

$$p(X \leq a) = p(X < a) = F(a) = \int_{-\infty}^{a} f(x)dx$$



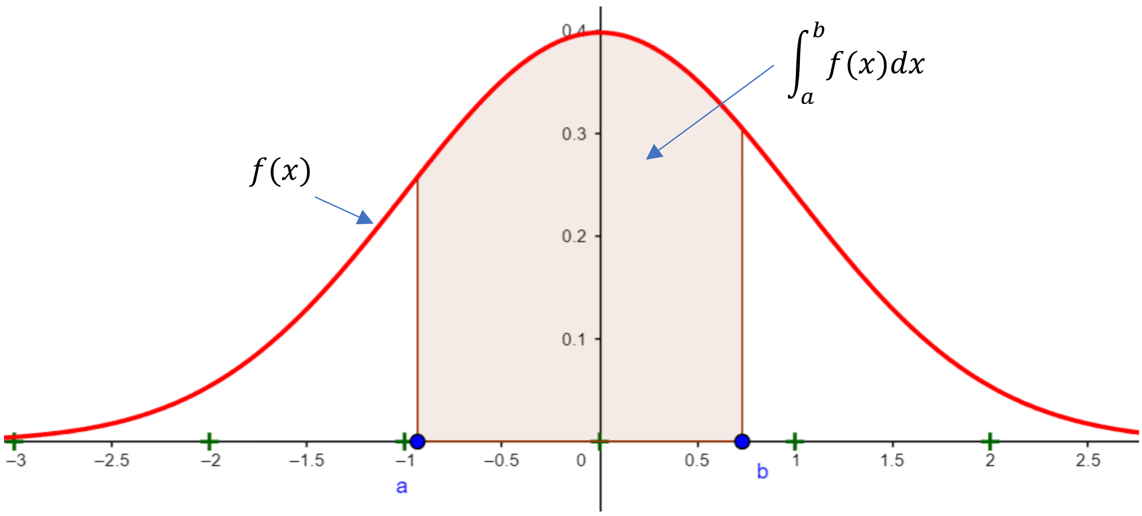
3. Para el cálculo de $p(X \geq \mathrm{b})$ se toma $1 - F(b)$:

$$p(X \geq b) = p(X > b) = 1 - F(b) = 1 - \int_b^{\infty} f(x)dx$$



4. Para el cálculo de $p(a \leq X \leq b)$ se evalúa F(x) en el intervalo [a,b]:

$$p(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(x)dx$$



Nótese que, para toda variable aleatoria CONTINUA:

$$p(a \leq X \leq b) = p(a \leq X < b) = p(a < X \leq b) = p(a < X < b)$$

1.6.5. INDEPENDENCIA DE VARIABLES ALEATORIAS CONTINUAS

Las variables aleatorias X e Y son independientes si se cumple:

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

Es decir, la MISMA definición de independencia que para variables DISCRETAS, pero contemplando que P(Z=z) = 0 en el caso de variables CONTINUAS.

1.6.6. ESPERANZA E(X) DE VARIABLES CONTINUAS

Se define como:

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

EJEMPLO

Se desea conocer la esperanza E(X) de una variable aleatoria cuya función de densidad es:

$$f(x)\begin{cases} 0 & si\ x \leq 2 \\ e^{2-x} & si\ x > 2 \end{cases}$$

Se desea calcular:

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

En este caso:

$$E(X) = \int_{-\infty}^{2} x \cdot 0\, dx + \int_{2}^{\infty} x \cdot e^{2-x} dx \underbrace{=}_{\substack{integrando \\ por\ partes}} 3$$

1.6.7. PROPIEDADES DE LA ESPERANZA E(X)

Las MISMAS propiedades de la esperanza de distribuciones DISCRETAS son válidas para distribuciones CONTINUAS:

1. La esperanza de una constante es esa misma constante.
2. La esperanza del producto de una variable aleatoria discreta X por una constante k escala en un factor k la esperanza E(X):

$$E(kX) = kE(X)$$

3. La esperanza de la suma de variables aleatorias es la suma de sus esperanzas:

$$E(X + Y) = E(X) + E(Y)$$

En el caso particular de una constante:

$$E(X + k) = E(X) + k$$

1.6.8. PROPIEDADES DE LA VARIANZA VAR(X)

Las MISMAS propiedades de la varianza de distribuciones DISCRETAS son válidas para distribuciones CONTINUAS:

1. La varianza de una constante es 0:

$$Var(X) = E(X^2) - E(X)^2 = k^2 - k^2 = 0$$

2. La varianza del producto de una variable aleatoria X por una constante k equivale al producto de la varianza de X por el CUADRADO de k:

$$Var(kX) = k^2 Var(X)$$

   De lo cual, para la desviación típica, se cumple:

$$\sigma(kX) = k\sigma(X)$$

3. La varianza de la suma de variables aleatorias INDEPENDIENTES equivale a la suma de sus varianzas:

$$Var(X + Y) = Var(X) + Var(Y)$$

   En el caso particular de una constante, nótese que Var(X) no varía:

$$Var(X + k) = Var(X)$$

1.7. Modelos de distribución continua

1.7.1. DISTRIBUCIÓN UNIFORME CONTINUA

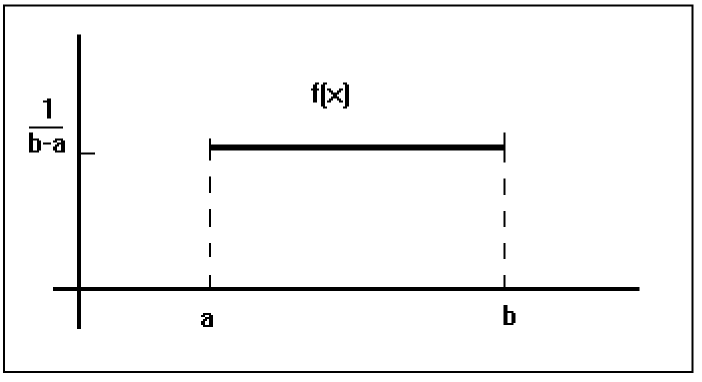
1.7.1.1. Concepto de distribución UNIFORME CONTINUA

Se denota como:

$$X \sim U(a, b)$$

Se define mediante la función de densidad:

$$f(x)\begin{cases} \dfrac{1}{b-a} & si\ a \leq x \leq b \\ 0 & en\ (-\infty, a) \cup (b, \infty) \end{cases}$$



Nótese que la distribución UNIFORME concede a todos los valores del intervalo en que está definida la MISMA probabilidad.

1.7.1.2. Cálculo de probabilidad

La probabilidad de la distribución uniforme verifica:

$$p(X > x_0) = \int_{x_0}^{\infty} f(x)dx \underbrace{=}_{\substack{Para\ x>b \\ f(x)\ es\ NULA}} \int_{x_0}^{b} \frac{1}{b-a} dx = \left[\frac{1}{b-a} x\right]_{x_0}^{b}$$

De igual modo:

$$p(X < x_0) = \int_{-\infty}^{x_0} f(x)dx \underbrace{=}_{\substack{Para\ x<a \\ f(x)\ es\ NULA}} \int_{a}^{x_0} \frac{1}{b-a} dx = \left[\frac{1}{b-a} x\right]_{a}^{x_0}$$

1.7.1.3. Esperanza de la distribución uniforme

La esperanza de la variable es:

$$E(X) = \frac{1}{2}(b+a)$$

Se demuestra fácilmente:

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx = \int_{a}^{b} x \cdot \frac{1}{b-a} dx = \frac{1}{b-a}\int_{a}^{b} xdx = \left| \frac{1}{b-a} \cdot \frac{1}{2}x^2 \right|_a^b$$

$$E(X) = \frac{1}{b-a} \cdot \frac{1}{2}b^2 - \frac{1}{b-a} \cdot \frac{1}{2}a^2 = \frac{1}{2(b-a)}(b^2-a^2) = \frac{1}{2(b-a)}(b+a)(b-a) = \frac{1}{2}(b+a)$$

1.7.1.4. Varianza de la distribución uniforme

La varianza de la variable es:

$$Var(X) = \frac{(b-a)^2}{12}$$

Se demuestra análogamente a lo anterior:

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

Donde:

$$E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx = \int_{a}^{b} x^2 \cdot \frac{1}{b-a} dx$$

Es decir:

$$E(X^2) = \left| \frac{1}{b-a} \cdot \frac{1}{3}x^3 \right|_a^b = \frac{1}{b-a} \cdot \frac{b^3}{3} - \frac{1}{b-a} \cdot \frac{a^3}{3} = \frac{1}{3}\frac{1}{b-a}(b^3-a^3)$$

De lo cual, el término $E(X^2)$ es:

$$E(X^2) = \frac{1}{3}\frac{1}{b-a}(b^3-a^3) = \frac{1}{3}\frac{1}{b-a}(b-a)(b^2+ba+a^2) = \frac{b^2+ba+a^2}{3}$$

Y el término $E(X)^2$:

$$E(X)^2 = \left[\frac{1}{2}(b+a)\right]^2 = \frac{1}{4}(b+a)^2$$

Entonces:

$$Var(X) = E(X^2) - E(X)^2 = \frac{b^2+ba+a^2}{3} - \frac{b^2+2ab+a^2}{4}$$

Operando:

$$Var(X) = \frac{1}{12}(4b^2+4ba+4a^2-3b^2-6ab-3a^2) = \frac{1}{12}(b^2-2ab+a^2)$$

$$Var(X) = \frac{1}{12}(b-a)^2$$

1.7.1.5. Función de distribución F(x) de la distribución uniforme

Una variable uniforme continua satisface la función de distribución F(x):

$$F(x)\begin{cases} 0 & si\ x < a \\ \dfrac{x-a}{b-a} & si\ a \leq x \leq b \\ 1 & si\ x > b \end{cases}$$

Nótese que, por definición de F(x), se tiene:

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

En este caso, la integral de la función de densidad f(x) desde el LÍMITE INFERIOR en que está definida sin ser nula, es decir, desde a, HASTA x (esto es la probabilidad ACUMULADA):

$$F(x) = \int_{-\infty}^{x} f(x)dx = \int_{a}^{x} \frac{1}{b-a}dx = \left|\frac{x}{b-a}\right|_{a}^{x} = \frac{x}{b-a} - \frac{a}{b-a} = \frac{x-a}{b-a}$$

1.7.1.6. Aproximación de la distribución uniforme a Normal (TCL)

El Teorema Central del Límite (TCL o TLC) demuestra que, para cualquier tamaño de muestra n > 30, una variable de cualquier distribución X incluida, en este caso, una distribución uniforme $X \sim U(a,b)$, se puede aproximar con un error despreciable a otra variable Y que sigue una Normal $Y \sim N(\mu, e_{\bar{x}})$, cuya esperanza es la esperanza μ de X y desviación típica es el error estándar de la variable X, calculado como $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ si se conoce Var(x) o bien $s_{\bar{x}} = \frac{\hat{s}}{\sqrt{n}}$ si no se conoce Var(X).

EJEMPLO

La concentración de un contaminante se distribuye de forma uniforme en el intervalo de 0 a 20 ppm (partes por millón). Una concentración se considera tóxica a partir de 8 ppm.

a) Se desea calcular qué probabilidad hay de que al tomar una muestra su concentración resulte tóxica.

b) Calcular la concentración media y la varianza.

c) Se desea calcular la probabilidad de que la concentración sea 10 ppm.

Se tiene:

X = "concentración de un contaminante en ppm".

$$X \sim U(0{,}20)$$

La función de densidad para una distribución UNIFORME es:

$$f(x)\begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & (-\infty, a) \cup (b, \infty) \end{cases} \rightarrow f(x)\begin{cases} \dfrac{1}{20-0} & 0 \le x \le 20 \\ 0 & (-\infty, a) \cup (b, \infty) \end{cases}$$

Por tanto, se puede calcular la función de distribución F(x):

$$F(x) = \begin{cases} 0 & x < 0 \\ \displaystyle\int_0^{20} \frac{1}{20} dx = \frac{1}{20} x & 0 \le x \le 20 \\ 1 & x > 20 \end{cases}$$

a) La probabilidad de toxicidad (más de 8 ppm) en una muestra es:

$$p(X > 8) = \int_8^{20} \frac{1}{20} dx = \left[\frac{1}{20} x\right]_8^{20} = \frac{1}{20}(20-8) = \frac{12}{20} = \frac{3}{5}$$

O bien, aplicando la expresión de la función de distribución F(x):

$$p(X \ge b) = p(X > b) = 1 - F(b) = 1 - \frac{1}{20} \cdot x$$

En este caso:

$$p(X > 8) = 1 - p(X < 8) = 1 - \frac{1}{20} \cdot 8 = \frac{20}{20} - \frac{8}{20} = \frac{12}{20} = \frac{3}{5}$$

b) La concentración media corresponde con esperanza de la variable:

$$E(x) = \frac{0+20}{2} = 10\ ppm$$

La varianza de la variable es:

$$Var(x) = \frac{(20-0)^2}{12} = \frac{400}{12} = \frac{100}{3}\ ppm$$

c) La variable X sigue una distribución CONTINUA. Entonces, la probabilidad acumulada en la función de distribución F(x) en un VALOR PUNTUAL (y no un intervalo) es NULA:

$$p(X = 10) = \int_{10}^{10} \frac{1}{20} dx = \left[\frac{1}{20} x\right]_{10}^{10} = \frac{1}{20}(10-10) = 0$$

EJEMPLO

El volumen de ventas se un negocio se distribuye de forma UNIFORME entre 38.000 y 120.000 €. Se desea calcular la probabilidad de facturar más de 100.000 €, la esperanza de la variable y su varianza.

Se tiene $X \sim U(38{,}120)$ (se expresa en miles de € para mejor lectura).

Por tanto:

$$f(x)\begin{cases} \dfrac{1}{120-38} & 38 \leq x \leq 120 \\ 0 & (-\infty, 38) \cup (120, \infty) \end{cases}$$

La probabilidad pedida es:

$$p(X > 100) = \int_{100}^{120} \frac{1}{120-38} dx = \left[\frac{x}{82}\right]_{100}^{120} = \frac{120}{82} - \frac{100}{82} = 0.2439$$

La esperanza de la variable es:

$$E(X) = \frac{120+38}{2} = 79$$

La varianza de la variable es:

$$Var(X) = \frac{(120-38)^2}{12} = 560.\hat{3}$$

1.7.2. DISTRIBUCIÓN EXPONENCIAL

1.7.2.1. Concepto

Está relacionada con la distribución de Poisson en cuanto a que se suele aplicar a modelar la PROBABILIDAD DEL TIEMPO entre sucesos CONSECUTIVOS. Se denota como:

$$X \sim Exp(\lambda)$$

Donde el parámetro λ corresponde al INVERSO DE LA ESPERANZA de X. Es decir:

$$\lambda = \frac{1}{E(X)}$$

O sea:

$$X \sim Exp\left(\frac{1}{E(X)}\right)$$

Nótese que, en la distribución de Poisson, λ se relaciona con la esperanza DIRECTAMENTE:

$$Si\ X \sim Po(\lambda) \rightarrow \lambda = E(X)$$

En cambio, en la distribución exponencial, λ se relaciona con el INVERSO de la esperanza E(X):

$$Si\ X \sim Exp(\lambda) \rightarrow \lambda = \frac{1}{E(X)}$$

1.7.2.2. Función de densidad f(x)

Se define mediante la función de densidad:

$$f(x) \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

1.7.2.3. Esperanza de la distribución Exponencial

La esperanza de la variable es $\frac{1}{\lambda}$, es decir, el inverso del parámetro:

$$E(x) = \frac{1}{\lambda}$$

1.7.2.4. Varianza de la distribución Exponencial

La varianza de la variable es:

$$Var(x) = \frac{1}{\lambda^2}$$

1.7.2.5. Función de distribución Exponencial

La función de distribución F(x) se tiene integrando la función de densidad f(x):

$$F(x) = \int_0^x \lambda e^{-\lambda x}\, dx = -\int_0^x -\lambda e^{-\lambda x}\, dx = \left[-e^{-\lambda x}\right]_0^x = \left(-e^{-\lambda \cdot x}\right) - \left(-e^{-\lambda \cdot 0}\right)$$

De lo cual: $F(x) = -e^{-\lambda x} - (-e^0) = -e^{-\lambda x} + 1 = 1 - e^{-\lambda x}$

Es decir:

$$F(x) = 1 - e^{-\lambda x}$$

1.7.2.6. Cálculo de probabilidad de distribución Exponencial

La probabilidad en un intervalo (a,b) para a, b > 0 es:

$$p(a < X < b) = \int_a^b f(x)dx = \int_{x_0}^b \lambda e^{-\lambda x}\, dx = -\int_{x_0}^b -\lambda e^{-\lambda x}\, dx = \left[-e^{-\lambda x}\right]_a^b$$

Es decir:

$$p(a < X < b) = \left(-e^{-\lambda b}\right) - \left(-e^{-\lambda a}\right) = -e^{-\lambda b} + e^{-\lambda a}$$

EJEMPLO

Sea una variable Y sigue una distribución de Poisson que modela el número de bostezos por unidad de tiempo en una sesión de estadística (frecuencia de los bostezos). En ese caso, la distribución toma como parámetro λ la ESPERANZA E(X):

$$Y \sim Po\left(\underbrace{\lambda}_{E(X)}\right)$$

Se puede definir otra variable Z que modela el TIEMPO (en minutos) que pasa entre un bostezo y el siguiente.

Si esa variable Z sigue una distribución EXPONENCIAL, toma como parámetro λ, pero en el caso de la EXPONENCIAL, λ NO ES LA ESPERANZA, SINO EL INVERSO DE LA ESPERANZA 1/E(X):

$$Z \sim Exp\left(\underbrace{\lambda}_{\frac{1}{E(X)}}\right)$$

Si la exponencial Z tiene como parámetro $\lambda = 2$, es decir:

$$Z \sim Exp(2)$$

Su esperanza es:

$$E(Z) = \frac{1}{\lambda} = \frac{1}{2}$$

Lo cual significa que entre bostezos hay, en promedio, un tiempo de medio minuto.

EJEMPLO

La esperanza de vida de un componente electrónico sigue una distribución EXPONENCIAL con media 8 años (nótese que **NO DICE** "parámetro 8").
Se pide calcular la probabilidad de que la vida de una muestra de ese componente esté entre 3 y 12 años.

Nótese que el parámetro λ es el inverso de la media. Por tanto, se tiene $\lambda = \frac{1}{8}$ en la definición de la variable:

$$X \sim E\left(\frac{1}{8}\right)$$

Es decir, si la media es 8, el parámetro es su inverso 1/8.

Entonces se cumple que la esperanza, que corresponde con la edad media del componente, es el INVERSO DEL PARÁMETRO λ, es decir:

$$E(x) = \frac{1}{\lambda}$$

En este caso, con $\lambda = \frac{1}{8}$ se tiene una esperanza:

$$E(x) = \frac{1}{\lambda} = \frac{1}{\frac{1}{8}} = 8$$

La varianza de la variable es:

$$Var(x) = \frac{1}{\left(\frac{1}{8}\right)^2} = \frac{1}{\frac{1}{64}} = 6$$

La función de densidad es:

$$f(x) \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & (-\infty, a) \cup (b, \infty) \end{cases}$$

En este caso:

$$f(x) \begin{cases} \frac{1}{8} e^{-\frac{1}{8}x} & x > 0 \\ 0 & (-\infty, a) \cup (b, \infty) \end{cases}$$

Integrando y evaluando en el intervalo [3,12], se calcula la probabilidad deseada:

$$p(3 < X < 12) = \int_3^{12} \frac{1}{8} e^{-\frac{1}{8}x} dx = -\int_3^{12} -\frac{1}{8} e^{-\frac{1}{8}x} dx = \left[-e^{-\frac{1}{8}x}\right]_3^{12} = \left(-e^{-\frac{12}{8}}\right) - \left(-e^{-\frac{3}{8}}\right) = 0.4641$$

EJEMPLO

El tiempo que toma hacer la revisión de un avión toma, aproximadamente, una distribución exponencial, con una media de 22 minutos.

a) Calcular la probabilidad con que una revisión dura menos de 10 minutos.

b) El coste de la revisión es de 200€ por cada media hora o fracción de media hora. ¿Cuál es la probabilidad de que una revisión cueste 400€?

c) Para realizar una programación de las revisiones de motor, ¿cuánto tiempo debe asignarse a cada revisión para que la probabilidad de cualquier tiempo de revisión mayor que el tiempo asignado sea de 0.1?

Se tiene:
X = "tiempo de revisión del motor en minutos".
Y se sabe:

$$X \sim E(\lambda)$$

Donde λ es el inverso de la media. Entonces:

$$X \sim E\left(\frac{1}{22}\right)$$

La función de densidad que sigue una distribución exponencial es:

$$f(x)\begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

En este caso:

$$f(x)\begin{cases} \frac{1}{22} e^{-\frac{x}{22}} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

La función de distribución F(x) se consigue integrando la función de densidad f(x) entre 0 y x:

$$F(x) = \int_0^x \frac{1}{22} e^{-\frac{x}{22}} dx = -\int_0^x -\frac{1}{22} e^{-\frac{x}{22}}\, dx = \left[-e^{-\frac{x}{22}}\right]_0^x = \left(-e^{-\frac{x}{22}}\right) - \left(-e^{-\frac{0}{22}}\right) = -e^{-\frac{x}{22}} + 1$$

Por tanto, se escribe F(x) como:

$$F(x)\begin{cases} 0 & x < 0 \\ 1 - e^{-\frac{x}{22}} & x \geq 0 \end{cases}$$

a) Se pide P(X<10):

Integrando la función de densidad:

$$P(x<10)=\int_0^{10}\frac{1}{22}e^{-\frac{x}{22}}dx=\left[-e^{-\frac{x}{22}}\right]_0^{10}=\left(-e^{-\frac{10}{22}}\right)-\left(-e^{-\frac{0}{22}}\right)=-e^{-\frac{10}{22}}+1=0.3652$$

Alterativamente, usando la función de distribución F(10):

$$P(x<10)=F(10)=1-e^{-\frac{10}{22}}=0.3652$$

b) Un coste de 400€ representa una duración de entre 30 minutos y UN SEGUNDO (se ingresa en el segundo tramo de 30') y hasta 60 minutos (en el minuto 61 ya costaría 600). Por tanto, se desea calcular:

$$P(30<x\leq 60)=\int_{30}^{60}\frac{1}{22}e^{-\frac{x}{22}}\,dx=\left[-e^{-\frac{x}{22}}\right]_{30}^{60}=\left(-e^{-\frac{60}{22}}\right)-\left(-e^{-\frac{30}{22}}\right)=0.1903$$

Alternativamente, aplicando F(x) directamente:

$$P(30<x\leq 60)=F(60)-F(30)=\left(1-e^{-\frac{60}{22}}\right)-\left(1-e^{-\frac{30}{22}}\right)=0.1903$$

c) Considerando k = "Tiempo asignado a cada revisión", se desea averiguar qué valor de k cumple:

$$P(X>k)=0.1$$

Es decir:

$$P(X>k)=\int_k^{\infty}\frac{1}{22}e^{-\frac{x}{22}}\,dx=\left[-e^{-\frac{x}{22}}\right]_k^{\infty}=\left(-e^{-\frac{\infty}{22}}\right)-\left(-e^{-\frac{k}{22}}\right)=-\frac{1}{e^{\frac{\infty}{22}}}+e^{-\frac{k}{22}}=0.1$$

Por tanto:

$$-\frac{1}{e^{\frac{\infty}{22}}}+e^{-\frac{k}{22}}=-\frac{1}{\infty}+e^{-\frac{k}{22}}=0+e^{-\frac{k}{22}}=0.1$$

De lo cual:

$$e^{-\frac{k}{22}}=0.1$$

Aplicando Ln en ambos miembros:

$$Ln\left(e^{-\frac{k}{22}}\right)=Ln(0.1)$$

Se alcanza:

$$-\frac{k}{22}=-2.3025$$

De lo cual: $k=22\cdot 2.3025=50.65\,minutos$

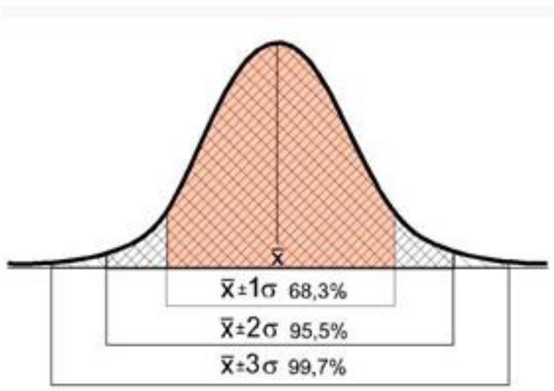
1.7.3. DISTRIBUCIÓN NORMAL

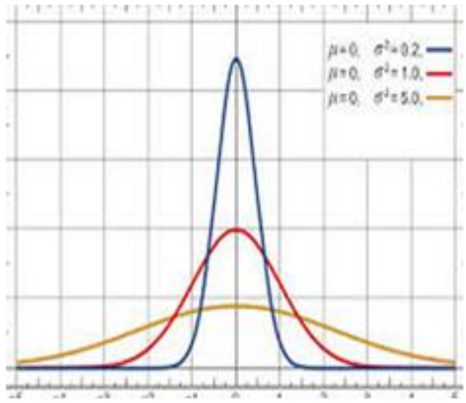
1.7.3.1. Concepto de distribución normal

Fue definida por Laplace y De Moivre a mediados del siglo XVIII y, de forma independiente, por Gauss (es la tópica campana de Gauss).
Se aplica al estudio de características biológicas, errores de fabricación, fenómenos naturales... su uso es MUY AMPLIO. Se denota por:

$$X \sim N(\mu, \sigma^2)$$

Pero esta expresión está en función de 2 parámetros (μ, σ), lo cual conlleva que EXISTEN INFINITAS distribuciones normales según los pares (μ, σ) que las definan.





ES FUNDAMENTAL ENTENDER QUE ES SIMÉTRICA.

1.7.3.2. Función de densidad de una distribución normal

La distribución normal se define por una función de densidad de la forma:

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

1.7.3.3. Esperanza de una distribución normal

Toma como esperanza el valor de la media μ:

$$E(X) = \mu$$

El proceso de demostración está aquí.

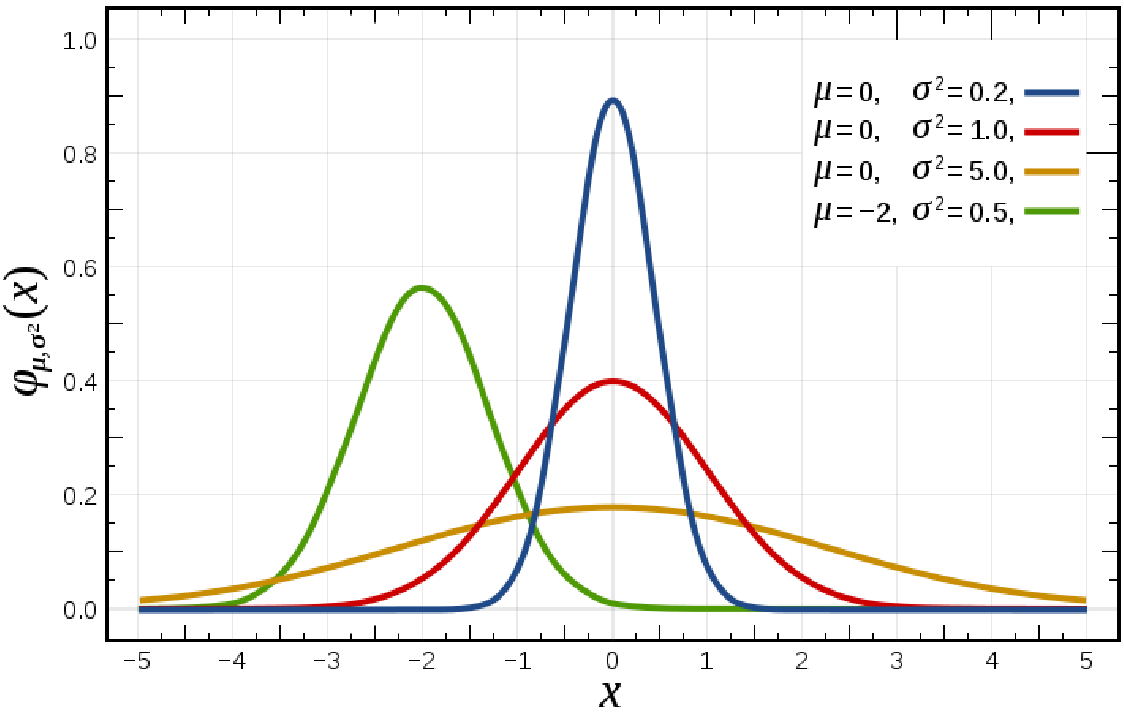
1.7.3.4. Varianza de una distribución normal

Toma como varianza el valor σ^2:

$$Var(X) = \sigma^2$$

El proceso de demostración está aquí.

1.7.3.5. Interpretación geométrica de los parámetros μ y σ



El valor de μ localiza el CENTRO de la distribución, por tanto, variar μ desplaza horizontalmente la distribución.

La desviación típica σ es una medida de la dispersión de lo datos y, por tanto, se observa una relación de PROPORCIONALIDAD entre la magnitud de σ y la ANCHURA de la distribución.

Los 2 valores de abscisa que corresponden a los PUNTOS DE INFLEXIÓN de la gráfica corresponden con $-\sigma$ y σ respectivamente. Es decir:

$$f(x) es \begin{cases} convexa \cup & en\ (-\infty, -\sigma) \\ cóncava \cap & en\ (-\sigma, +\sigma) \\ convexa \cup & en\ (+\sigma, +\infty) \end{cases}$$

Dispersión	**Forma**
Baja	ESTRECHA y PUNTIAGUDA
Alta	ANCHA y NO PUNTIAGUDA

1.7.3.6. Función de distribución de una normal

Su función de distribución F(x) no es sencilla, de modo que se expresa:

$$F(x) = \mathrm{P}(\mathrm{X} \leq \mathrm{x_0}) = \int_{-\infty}^{x_0} \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

No obstante, no suele usarse para el cálculo de probabilidad en distribuciones normales, sino que se recurre a la variable tipificada normal Z y a la tabla Z, que refleja los valores de su función de distribución $\Phi(x)$.

1.7.3.7. Suma de variables normales

Si se tienen 2 variables aleatorias INDEPENDIENTES que siguen distribuciones normales cuyas leyes son:

$$X \sim N(\mu_X, \sigma_X)$$
$$Y \sim N(\mu_Y, \sigma_Y)$$

Su suma (o su resta) es una variable V que sigue una ley normal:

$$X \pm Y = V \sim N(\mu_X \pm \mu_Y, \sigma_X \pm \sigma_Y)$$

Esto es extensible a n variables aleatorias INDEPENDIENTES que sigan distribuciones normales:

$$\sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma)$$

1.7.3.8. Transformación lineal de variables normales

Se cumple, para cualquier variable normal $X \sim N(\mu_X, \sigma_X)$ la transformación lineal:

$$a \cdot X + b \sim N\left(\underbrace{a \cdot \mu_X + b}_{\substack{nueva \\ media}}, \underbrace{a \cdot \sigma_X}_{\substack{nueva \\ desviación}}\right)$$

Es decir:
- La media experimenta el cambio de escala y de origen.
- La desviación típica experimenta el cambio de escala.

Nótese que es una combinación de cambio de origen y de escala aplicados a la media y a la desviación estándar.

1.7.4. DISTRIBUCIÓN NORMAL ESTÁNDAR Z Y TABLA Z

1.7.4.1. Concepto de normal estándar Z

Como trabajar con F(x) de distribuciones normales para el cálculo de probabilidad es muy complejo, se recurre a:

1. La tipificación o estandarización para reducir los cálculos de probabilidad de cualquier normal $X \sim N(\mu, \sigma)$ a una normal denominada Z estándar que sigue $Z \sim N(0{,}1)$.
   El proceso de tipificación se detalla a continuación, pero antes se describe la aplicación de la tabla Z al cálculo de probabilidad.

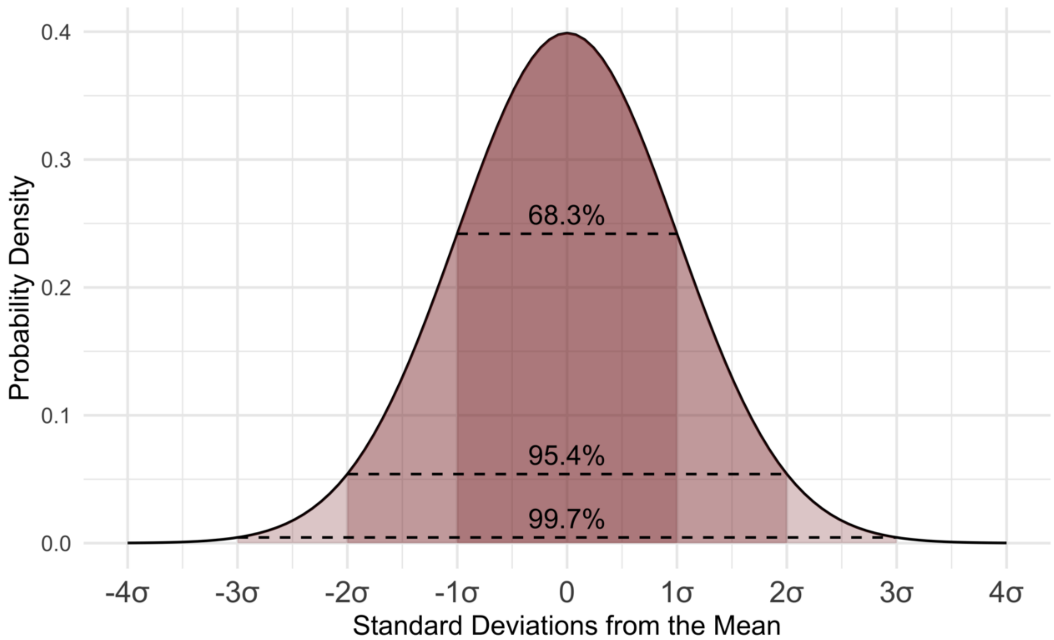
2. La tabulación de los resultados de la función de distribución de Z denotada por $\Phi(x)$ en la tabla denominada tabla Z.

La distribución normal estándar $Z \sim N(0{,}1)$ es el resultado de un proceso de TIPIFICACIÓN o ESTANDARIZACIÓN a partir de una variable normal $X \sim N(\mu, \sigma)$.

De entre las combinaciones posibles de μ y σ para definir una normal, se toma como arquetipo para calcular probabilidades la distribución de probabilidad de la normal con μ = 0 y $\sigma = 1$. Esa es la normal estándar Z.

Esto PARTICULARIZA la distribución normal X en una variable Z en concreto:

$$X \sim N(\mu, \sigma) \rightarrow Z \sim N(0{,}1)$$



Nótese que según la tabla Z, $P(X \leq 0) = 0.5$ (intersección de fila 0.0 y columna 0.00). Es decir, está centrada en 0.

1.7.4.2. Tipificación o ESTANDARIZACIÓN de variables normales: Relación entre normal X y normal estándar Z

Se cumple la relación:

$$Z = \frac{X-\mu}{\sigma}$$

Y, por tanto, también se da:

$$X = \mu + \sigma Z$$

NO ES IMPRESCINDIBLE LA DEMOSTRACIÓN QUE SIGUE.

Esta relación se demuestra:

→ Para esta demostración, por claridad, se usa:

- $F_x(x)$ es la función de distribución de X.
- $F_z(z)$ es la función de distribución de Z estándar.

Partiendo de la definición de la función de distribución de Z:

$$F_z(z) = P(Z \leq \mathrm{z})$$

Se considera la relación a verificar $Z = \frac{X-\mu}{\sigma}$:

$$F_z(z) = P(Z \leq \mathrm{z}) = P\left(\frac{X-\mu}{\sigma} \leq z\right)$$

Operando en la desigualdad $\frac{X-\mu}{\sigma} \leq z$, multiplicando ambos miembros por σ y sumando μ, se obtiene:

$$F_z(z) = P(Z \leq \mathrm{z}) = P\left(\frac{X-\mu}{\sigma} \leq z\right) = P(\mathrm{X} \leq \mu + \sigma z)$$

Lo cual es, por definición, un valor de la función de distribución de X:

$$P(\mathrm{X} \leq \mu + \sigma z) = F_x(\mu + \sigma z)$$

Entonces:

$$F_z(z) = P(\mathrm{X} \leq \mu + \sigma Z) = F_x(\mu + \sigma z)$$

Derivando respecto σ en ambos miembros, se expresa en términos de la función de densidad:

$$f_z(z) = f_x(\mu + \sigma z) \cdot \sigma$$

Es decir:

$$f_z(z) = \underbrace{\frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{((\mu+\sigma z)-\mu)^2}{2\sigma^2}}}_{f_x(\mu+\sigma z)} \cdot \sigma$$

Lo cual devuelve, efectivamente, la función de densidad de la variable normal estándar f(z):

$$f_z(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(\mu+\sigma z-\mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}$$

Análogamente, se demuestra la simetría de la relación.

Esta relación tiene como CONSECUENCIA FUNDAMENTAL:

$$P(\mathrm{a} \le \mathrm{X} \le b) = P\left(\frac{\mathrm{a}-\mu}{\sigma} \le \underbrace{\frac{\mathrm{X}-\mu}{\sigma}}_{Z} \le \frac{b-\mu}{\sigma}\right)$$

Entonces, se concluye:

$$P(\mathrm{a} \le \mathrm{X} \le b) = P\left(\frac{\mathrm{a}-\mu}{\sigma} \le \mathrm{Z} \le \frac{b-\mu}{\sigma}\right)$$

O lo que es lo mismo:

$$P(X \le k) = P\left(Z \le \frac{k-\mu}{\sigma}\right)$$

EJEMPLO

Se tiene una variable $X \sim N(5,2)$.
Se desea calcular $P(X \le 8)$, $P(X \le 3)$ y $P(3 \le X \le 8)$.

Para ello, primero se TIPIFICA a una variable $Z \sim N(0,1)$:

$$P(X \le 8) \underbrace{=}_{Tipificación} P\left(z \le \frac{8-5}{2}\right) = P\left(z \le \frac{3}{2}\right) = P(z \le 1.5) \underbrace{=}_{Tabla\ Z} 0.9932$$

$$P(X \le 3) \underbrace{=}_{Tipificación} P\left(z \le \frac{3-5}{2}\right) = P(z \le -1) \underbrace{=}_{\substack{\text{se gira la cola} \\ \text{por simetría}}} P(z > 1)$$

$$P(X \le 3) = P(z > 1) \underbrace{=}_{Por\ complementariedad} 1 - P(z \le 1) \underbrace{=}_{Tabla\ Z} 1 - 0.8413$$

$$P(3 \le X \le 8) = \underbrace{P(X \le 8)}_{\substack{Extremo\ superior \\ de\ intervalo}} - \underbrace{P(X \le 3)}_{\substack{Extremo\ inferior \\ de\ intervalo}} = 0.9932 - (1 - 0.8413) = 0.7745$$

EJEMPLO

Se tiene una variable X con media 50 y desviación 10.
Se desea calcular la probabilidad de que X adopte un valor entre 45 y 62.

Se tiene $X \sim N(50{,}10)$
Estandarizando:

$$P(45 < X < 62) = P\left(\frac{45-\mu}{\sigma} < Z < \frac{62-\mu}{\sigma}\right) = P\left(\frac{45-50}{10} < Z < \frac{62-50}{10}\right)$$

Es decir:

$$P(-0.5 < Z < 1.2) = \underbrace{P(Z < 1.2)}_{\substack{Tabla\ Z: \\ \text{fila 1.2 columna 0.00}}} - P(Z < -0.5) = 0.8849 - \underbrace{P(Z > 0.5)}_{por\ simetría}$$

$$P(-0.5 < z < 1.2) = 0.8849 - \underbrace{\big(1 - P(z < 0.5)\big)}_{por\ complementariedad} = 0.8849 - 1 + \underbrace{P(z < 0.5)}_{\substack{Tabla\ Z: \\ \text{fila 0.5 columna 0.00}}}$$

$$P(-0.5 < z < 1.2) = 0.8849 - 1 + 0.6915 = 0.5764$$

1.7.4.3. Función de densidad de la ley normal estándar

La ley normal posee una propiedad EXTREMADAMENTE ÚTIL: permite reducir el análisis de la probabilidad de cualquier distribución normal al análisis de una normal Z, es decir, una normal ESTÁNDAR.

Además, permite hacer comparables los resultados de probabilidad de diversas distribuciones normales.

Se considera una variable $X \sim N(\mu, \sigma)$. La probabilidad de X no corresponde con tabla de la distribución Z (a menos que $\mu = 0$ y $\sigma = 1$, es decir, que YA SEA ESTÁNDAR).

Para poder aplicar la tabla Z para el cálculo de probabilidad, se debe TIPIFICAR X, es decir ESTANDARIZARLA:

$$X \sim N(\mu, \sigma) \rightarrow Z \sim N(0{,}1)$$

La función de densidad de la normal Z es:

$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}$$

1.7.4.4. Función de distribución de normal estándar $\Phi(x)$

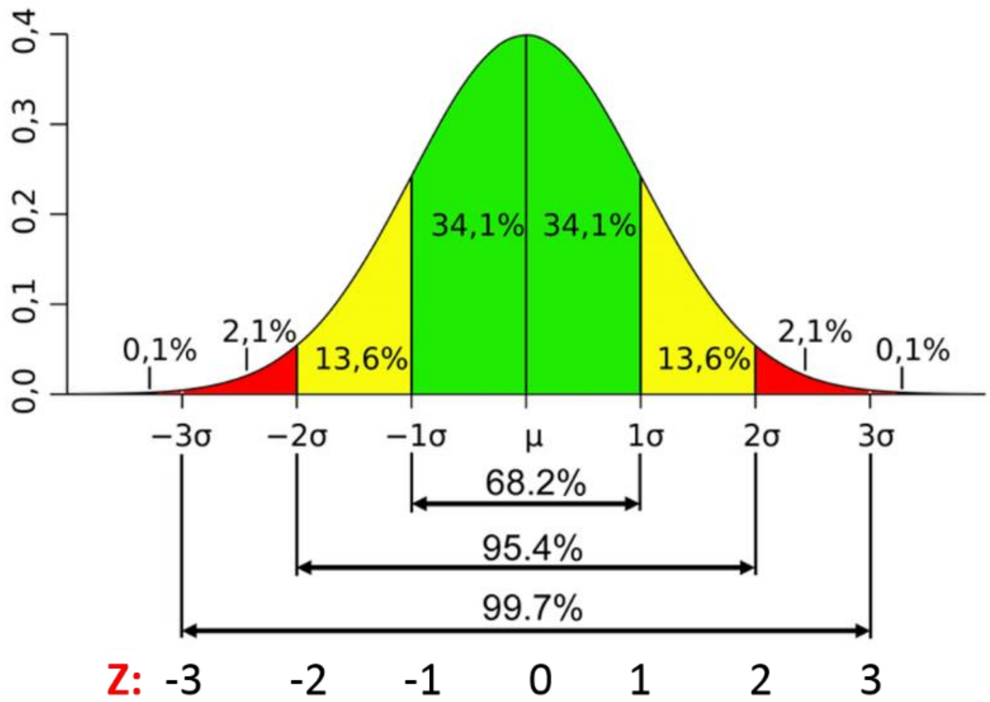
Se cumple:

$$\Phi(x) = P(Z \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}} dz$$

1.7.4.5. Probabilidad en la normal estándar

La normal estándar Z verifica:

- El 68% de la probabilidad se encuentra en el intervalo $(\mu - \sigma, \mu + \sigma)$
- El 95% de la probabilidad se encuentra en el intervalo $(\mu - 2\sigma, \mu + 2\sigma)$
- El 99.7% de la probabilidad se encuentra en el intervalo $(\mu - 3\sigma, \mu + 3\sigma)$

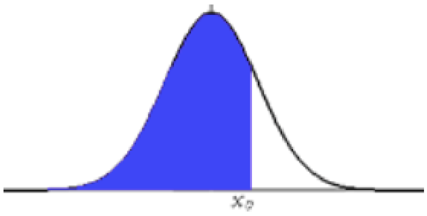


1.7.4.6. Tabla estándar Z

La tabla Z recoge los resultados de la función de distribución $\Phi(x)$ para cada valor de abscisa x_0, es decir, el área cerrada bajo la curva que representa la probabilidad con que la variable Z adopta un valor inferior o igual a x_0. Se lee:

- Las FILAS corresponden a las DÉCIMAS del valor x_0.
- Las COLUMNAS corresponden a las DÉCIMAS del valor x_0.
- La INTERSECCIÓN de filas y columnas corresponde a la PROBABILIDAD.

El valor que ocupa las filas y columnas, es decir, el valor de la abscisa es el ESTADÍSTICO z. Por ejemplo, el estadístico z = 0.01 corresponde a una probabilidad de 0.5040.



$P[X \leq x_0]$

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7703	0,7734	0,7704	0,7793	0,7823	0,7652
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	0,8364	0,8389
1,0	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	0,8930
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	0,9235	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	0,9485	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	0,9762	0,9767
2,0	0,9773	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	0,9934	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9865	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	0,9975	0,9975	0,9976	0,9977	0,9978	0,9978	0,9979	0,9980	0,9980	0,9981
2,9	0,9981	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0.9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999

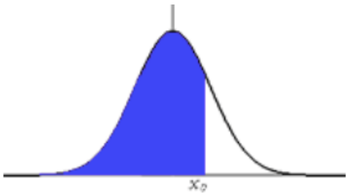
1.7.5. USO DE LA TABLA Z DE LA DISTRIBUCIÓN NORMAL (ESTADÍSTICO Z)

El cálculo de probabilidad se reduce al cálculo integral: el área a la izquierda del valor de abscisa x_0 corresponde a la probabilidad $P(Z \leq x_0)$ con que Z adopta un valor menor o igual a x_0.

La tabla contiene la probabilidad (el valor de F(x) asociado a cada x_0) con que Z adopta un valor inferior o igual a x_0.

Para leer la tabla, se descompone en décimas y centésimas el valor x_0 en filas y columnas respectivamente. La probabilidad es la intersección entre ellas.

EJEMPLO



$P[X \leq x_0]$

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7703	0,7734	0,7704	0,7793	0,7823	0,7652
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	0,8364	0,8389
1,0	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	0,8930
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	0,9235	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	0,9485	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	0,9762	0,9767
2,0	0,9773	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	0,9934	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9865	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	0,9975	0,9975	0,9976	0,9977	0,9978	0,9978	0,9979	0,9980	0,9980	0,9981
2,9	0,9981	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999

$P(Z \leq 0.36) = 0.6406$

Fila	0.3
Columna	0.06
Intersección	0.6406

$P(Z \leq 3.28) = 0.9995$

Fila	3.2
Columna	0.08
Intersección	0.9995

$P(Z \leq 0.025) = 0.501$

Fila	0.0
Columna	**.025??**

NO ESTÁ EN LA TABLA.

EL VALOR NO SE LEE, SE PROMEDIA entre las 2 columnas más próximas:

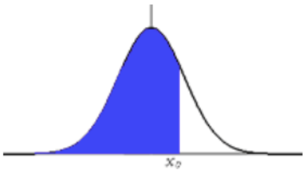
Columna	Intersección
0.02	0.5080
0.03	0.5120
Promedio:	**0.5010**

Nótese que promediar valores acarrea un ERROR, dado que la curva no es lineal.

La tabla Z también permite lo contrario, obtener el valor de abscisa x_0 asociado a que Z adopte una probabilidad dada:

EJEMPLO

¿Cuál es el valor mínimo x_0 con que Z se puede encontrar con un 83,6% de seguridad?



$P[X \leq x_0]$

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7703	0,7734	0,7704	0,7793	0,7823	0,7652
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	0,8364	0,8389
1,0	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	0,8930
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	0,9235	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	0,9485	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	0,9762	0,9767
2,0	0,9773	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	0,9934	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9865	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	0,9975	0,9975	0,9976	0,9977	0,9978	0,9978	0,9979	0,9980	0,9980	0,9981
2,9	0,9981	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0.9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999

El valor deseado está en el intervalo:

$$x_0 \in (0.97, 0.98)$$

Ya que la intersección 0.836 se encuentra en algún lugar entre las intersecciones de fila 0.9 con las columnas 0.07 y 0.08.

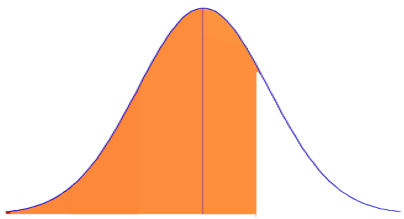
NO SE PUEDE ESTIMAR CON MÁS PRECISIÓN USANDO ESTA TABLA (interpolar linealmente, por ejemplo, acarrea un error, ya que no hay un comportamiento lineal.

1.7.5.1. Ejemplos de uso de la tabla z

a) Cálculo de $\boldsymbol{P(Z \leq x_0)}$

La probabilidad tabulada, es decir, el área tabulada, es en la distribución Z UNA COLA HACIA LA IZQUIERDA y, por tanto, contempla la probabilidad con que Z es MENOR O IGUAL que un valor x_0 dado: $P(Z \leq x_0)$.

EJEMPLO



$P(Z \leq 0.46) = 0.6772$

Fila	0.4
Columna	0.06
Intersección	0.6772

b) Cálculo de $\boldsymbol{P(Z \geq x_0)}$

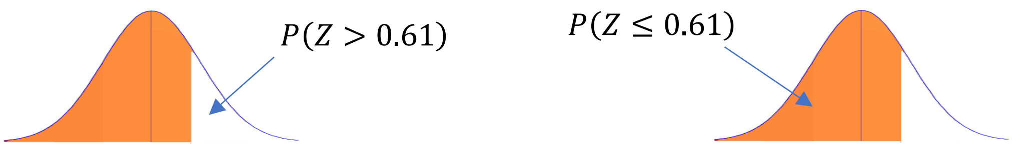
Complementariamente, la probabilidad con que X es MAYOR que un cierto x_0 se deduce fácilmente:

$$P(Z > x_0) = 1 - P(Z \leq x_0)$$

EJEMPLO

$$P(Z > 0.61) = 1 - P(Z < 0.61) = 1 - 0.7291$$

Fila	0.6
Columna	0.01
Intersección	0.7291



c) Cálculo de $\boldsymbol{P(Z \leq x_0)}$ con $x_0 < 0$

En la tabla Z no aparecen índices x_0 negativos.

Pero como **LA DISTRIBUCIÓN ES SIMÉTRICA**, se cumple:

$$P(Z \geq -x_0) = P(z < x_0)$$

O lo que es lo mismo:

$$P(Z \leq -x_0) = P(Z > x_0)$$

Es decir, girar la desigualdad y considerar el opuesto de la abscisa x_0 no varía el valor de probabilidad observado, ya que es **SIMÉTRICA**.

EJEMPLO

$P(Z < -1.41) = ?$

-1.41 no se encuentra en la tabla.

Basta con girar la cola para encontrar un área equivalente asociada a la abscisa z = 1.41:

$$P(Z < -1.41) = P(Z > 1.41)$$

Por último, sabiendo que la probabilidad total acumulada es 1:

$$P(Z > 1.41) = 1 - P(Z < 1.41) = 1 - 0.927 = 0.073$$

d) Cálculo de $\boldsymbol{P(x_0 \leq Z \leq x_1)}$ en un intervalo

Para calcular la probabilidad acumulada en un intervalo (a,b) basta con operar F(b) – F(a). Para ello, se deben considerar los argumentos de:

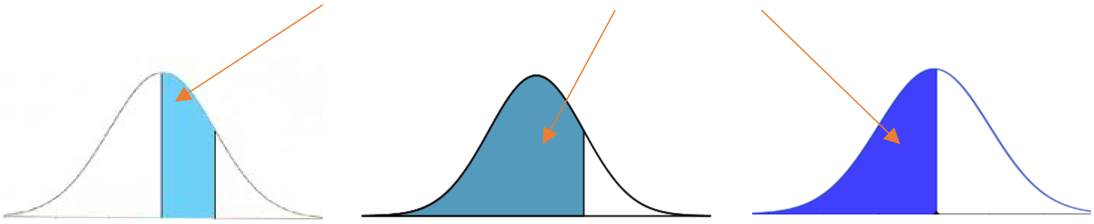
- SIMETRÍA $P(-\mathrm{a} \leq Z \leq +\mathrm{a})$
- COMPLEMENTARIEDAD $P(Z > a) = 1 - P(\mathrm{Z} \leq \mathrm{a})$

EJEMPLO

Se desea conocer $P(0 \leq Z < 1)$.

Se observa:

$$P(0 \leq Z < 1) = P(Z \leq 1) - P(Z \leq 0)$$



Entonces:

$$P(0 \leq Z < 1) = 0.8413 - 0.5 = 0.3413$$

EJEMPLO

Se desea conocer $P(-1 \leq Z \leq 1)$.

Para ello, se toma la probabilidad $P(Z \leq 1)$ y se le resta $P(Z \leq -1)$.

Es decir:

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

El primer valor $P(Z \leq 1)$ está en la tabla:

$$P(Z \leq 1) = 0.8413$$

Pero el segundo valor $P(Z \leq -1)$ no aparece. Se debe GIRAR la cola:

$$P(Z \leq -1) = 1 - P(Z > 1)$$

Por tanto:

$$P(-1 \leq Z \leq 1) = \underbrace{P(Z \leq 1)}_{Tabla\ Z} - \underbrace{P(Z \leq -1)}_{NO\ ESTÁ} = 0.8413 - \underbrace{P(Z > 1)}_{\substack{No\ hay\ valores \\ de\ P(Z>z_0)}}$$

Ahora bien, la tabla no refleja $P(Z > 1)$ sino solo $P(Z \leq 1)$. Entonces:

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

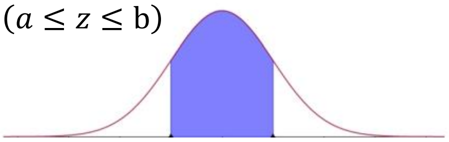
Y ahora sí se puede encontrar $P(z \leq 1)$ en la tabla. Es decir:

$$P(-1 \leq Z \leq 1) = 0.8413 - \underbrace{P(z > 1)}_{Por\ SIMETRÍA} = 0.8413 - \underbrace{(1 - P(z \leq 1))}_{\substack{por \\ COMPLEMENTARIEDAD}}$$

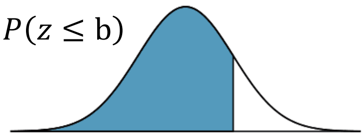
De lo cual:

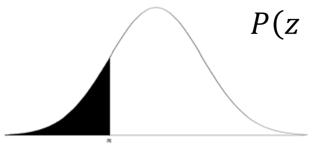
$$P(-1 \leq z < 1) = 0.8413 - (1 - 0.8413) = 0.6826$$

$P(a \leq z \leq \text{b})$



$P(z \leq \text{b})$





$P(z \leq \text{a})$

Estos 2 argumentos, SIMETRÍA y COMPLEMENTARIEDAD, son ESENCIALES en el trabajo con la tabla Z:

Argumento de SIMETRÍA $\underbrace{P(-a \leq z)}_{\substack{NO\ ESTÁ \\ en\ tabla\ Z\ :(}} = \underbrace{P(Z > +a)}_{\substack{NO\ ESTÁ \\ en\ tabla\ Z\ :(}}$

Argumento de COMPLEMENTARIEDAD $\underbrace{P(Z > a)}_{\substack{NO\ ESTÁ \\ en\ tabla\ Z\ :(}} = 1 - \underbrace{P(Z \leq a)}_{SÍ\ ESTÁ\ :D}$

Nótese que para un intervalo simétrico respecto 0, es decir, $P(-a \leq Z < a)$ se puede generalizar:

$$P(-a \leq Z < a) = 1 - 2 \cdot P(Z \leq a) = REGIÓN\ INCLUIDA$$