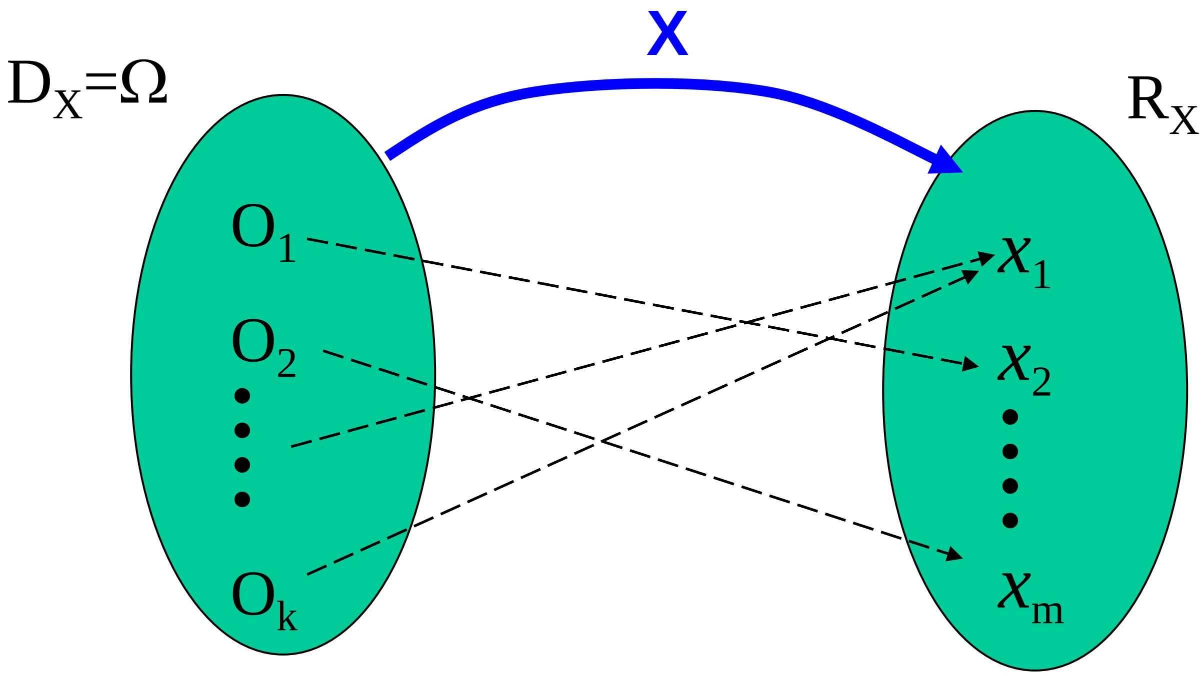
VARIABLES ALEATORIAS

Variable aleatoria

Es una función cuyo dominio es un espacio muestral, y cuyo rango es un subconjunto de lo números reales.



$D_X=\Omega=\{O_1, O_2, ..., O_k\}$

$R_X=\Omega=\{x_1, x_2, ..., x_m\}$

VARIABLES ALEATORIAS

Tipo de variables aleatorias

1. Variables aleatorias discretas

Son aquellas cuyos rangos forman conjuntos numerables. Pueden contener un número finito o infinito de elementos.

Ejemplos: Número de fallas por cada metro de roca.
Número de diaclasas por metro de roca.

2. Variables aleatorias continuas

Son aquellas cuyos rangos forman conjuntos no numerables.

Ejemplos: Concentración de Cobre en un taladro.

Concentración de Pb en una muestra de agua de laguna.

VARIABLES ALEATORIAS DISCRETAS.

Función de probabilidad

Sea X una variable aleatoria discreta que tiene como rango R_X ; luego, una función *f(x)* es llamada función de probabilidad de la variable aleatoria X si tiene como dominio a R_X, y como rango a un conjunto de números reales $P[X = x_i] = f(x_i)$ que cumplen con las siguientes condiciones.

1. $P[X=x_i] = f(x_i) \geq 0 \quad , \quad \forall\, x_i \in R_X$

2. $0 \leq f(x_i) \leq 1 \quad , \quad \forall\, x_i \in R_X$

3. $\sum_{x_i \in R_X} f(x_i) = 1$

VARIABLES ALEATORIAS DISCRETAS. **Ejemplo**

Analytics AoZ

Una ONG ha propuesto 10 nuevos proyectos, de los cuales 4 son proyectos educativos y 6 son proyectos sociales. Para realizar un análisis de factibilidad de los proyectos se eligen al azar 3 de ellos para que sean analizados la próxima semana.

Si se define X= Número de proyectos educativos que serán evaluados la próxima semana, halle la función de probabilidad de la variable X.

Exp. Aleatorio: Se eligen al azar 3 proyectos para ser analizados.

4 P.Educativos 6 P.Sociales

$\xrightarrow{\text{MSR}}$ 3

$$n(\Omega) = \binom{10}{3} = 120$$

VARIABLES ALEATORIAS DISCRETAS. Ejemplo

Ω={ (S_1,S_2,S_3), (S_1,S_2,E_3), (S_1,E_2,S_3), (E_1,S_2,S_3), (E_1,E_2,S_3), (E_1,S_2,E_3), (S_1,E_2,E_3), (E_1,E_2,E_3) },

X	0	1	1	1	2	2	2	3

$\Rightarrow R_X=\{0,1,2,3\}$

$f(0)=P[\ X=0\ ]=P[\ S_1,S_2,S_3\ ]=(6/10)(5/9)(4/8)=120/720=5/30$

$f(1)=P[\ X=1\ ]=P[\ S_1,S_2,E_3\ ]+P[\ S_1,E_2,S_3\ ]+P[\ E_1,S_2,S_3\ ]$

$=(6/10)(5/9)(4/8)+(6/10)(4/9)(5/8)+(4/10)(6/9)(5/8)=15/30$

$f(2)=P[\ X=2\ ]=P[\ S_1,E_2,E_3\ ]+P[\ E_1,S_2,E_3\ ]+P[\ E_1,E_2,S_3\ ]$

$=(6/10)(4/9)(3/8)+(4/10)(6/9)(3/8)+(4/10)(3/9)(6/8)=9/30$

$f(3)=P[\ X=3\ ]=P[\ E_1,E_2,E_3\ ]=(4/10)(3/9)(2/8)=24/720=1/30$

VARIABLES ALEATORIAS DISCRETAS. **Ejemplo**

Analytics AoZ

Función de probabilidad de la variable aleatoria X

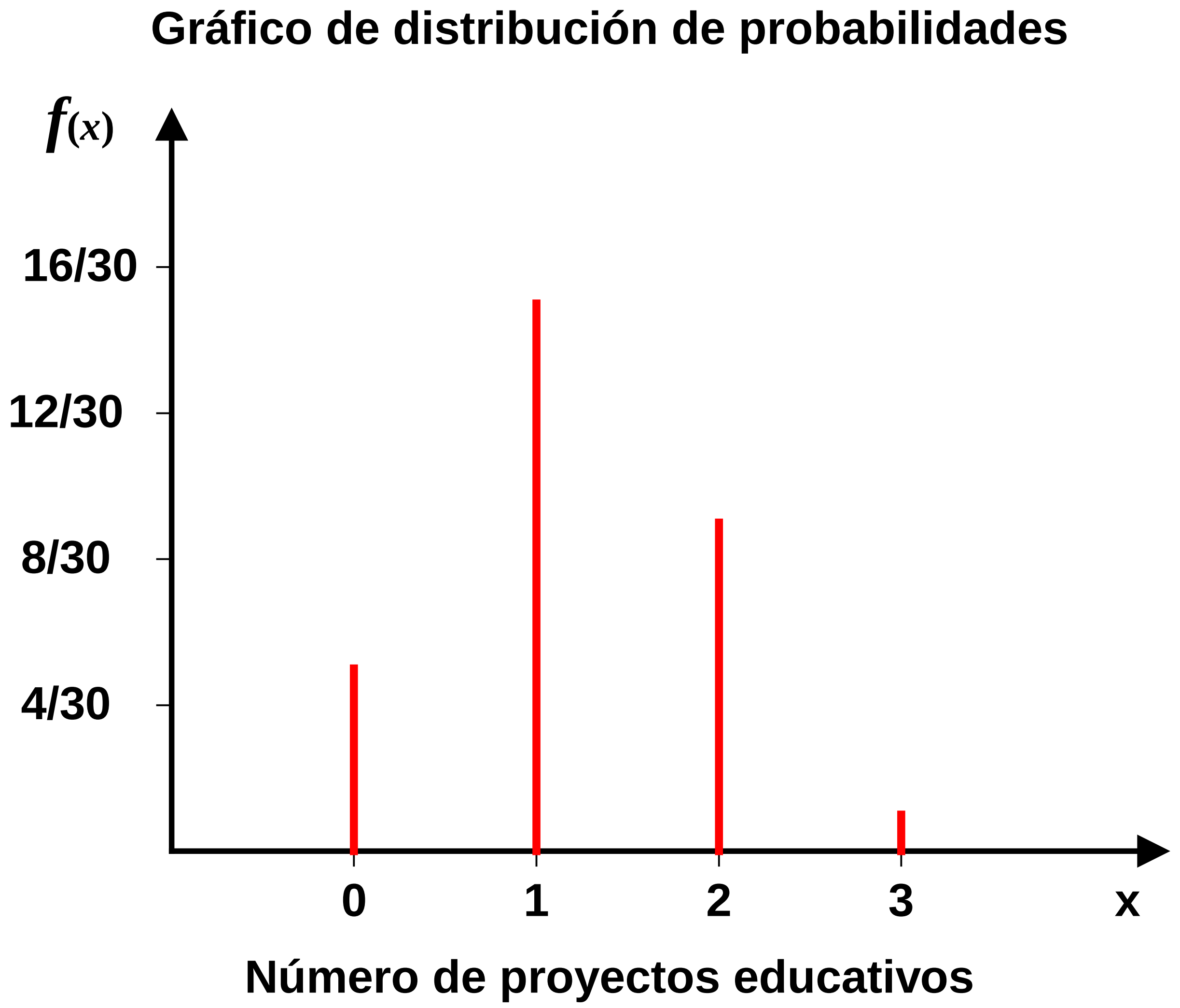
$$
\begin{aligned}
f(x) &= 5/30, & &\text{si } x = 0\\
&= 15/30, & &\text{si } x = 1\\
&= 9/30, & &\text{si } x = 2\\
&= 1/30, & &\text{si } x = 3\\
&= 0, & &\text{otros valores de } x
\end{aligned}
$$

Tabla de probabilidades

X	0	1	2	3
$f(x)$	5/30	15/30	9/30	1/30

VARIABLES ALEATORIAS DISCRETAS. Ejemplo

Analytics AoZ



VARIABLES ALEATORIAS DISCRETAS.

Función de probabilidad acumulativa

Sea X una variable aleatoria discreta que tiene como función de probabilidad *f(x)*; luego, la función de probabilidad acumulativa de la variable aleatoria X se define como:

$$F(x) = P[X \leq x] = \sum_{x_i \leq x} f(x_i)$$

VARIABLES ALEATORIAS DISCRETAS.

Analytics AoZ

Función de probabilidad acumulativa.

Propiedades

1. $F(x)=0, \quad \forall \ x<m, \quad donde \quad m=Min \quad x\in R_X$

2. $F(x)=1, \quad \forall \ x\geq M, \quad donde \quad \mathrm{M}=Max \quad x\in R_X$

3. $0\leq F(x)\leq 1, \quad \forall \ x\in R_X$

4. $\mathrm{F}(x)$ es una función no decreciente.

VARIABLES ALEATORIAS DISCRETAS. **Ejemplo**

Analytics AoZ

Si el número de proyectos educativos que son evaluados por semana es una variable aleatoria X que tiene como tabla de probabilidades:

X	0	1	2	3
$f(x)$	5/30	15/30	9/30	1/30

Halle la función de distribución acumulativa de X.

Puesto que $R_X=\{0,1,2,3\} \Rightarrow$ **m**=Min $x\in R_X$ =**0**, **M**=Max $x\in R_X$ =**3**,

$F(\boldsymbol{x})=P[\ X < \boldsymbol{x}\ ]=0 \quad \forall\ \boldsymbol{x} <0$

VARIABLES ALEATORIAS DISCRETAS. **Ejemplo**

X	0	1	2	3
$f(x)$	5/30	15/30	9/30	1/30

$R_X=\{0,1,2,3\} \Rightarrow$ **m**=Min $x\in R_X$ =**0**, **M**=Max $x\in R_X$ =**3**,

Para $0 \leq x < 1$

$F(0)=P[\ X \leq 0\ ]= P[\ X < 0\ ]+P[\ X = \mathbf{0}\ ]= 0 + f(\mathbf{0}) = 5/30$

Para $1 \leq x < 2$

$F(1)=P[\ X \leq 1]= P[\ X < 1\ ]+P[\ X = \mathbf{1}\ ]= 5/30 + f(\mathbf{1}) = 20/30$

Para $2 \leq x < 3$

$F(2)=P[\ X \leq 2\ ]= P[\ X < 2\ ]+P[\ X = \mathbf{2}\ ]= 20/30 + f(\mathbf{2}) = 29/30$

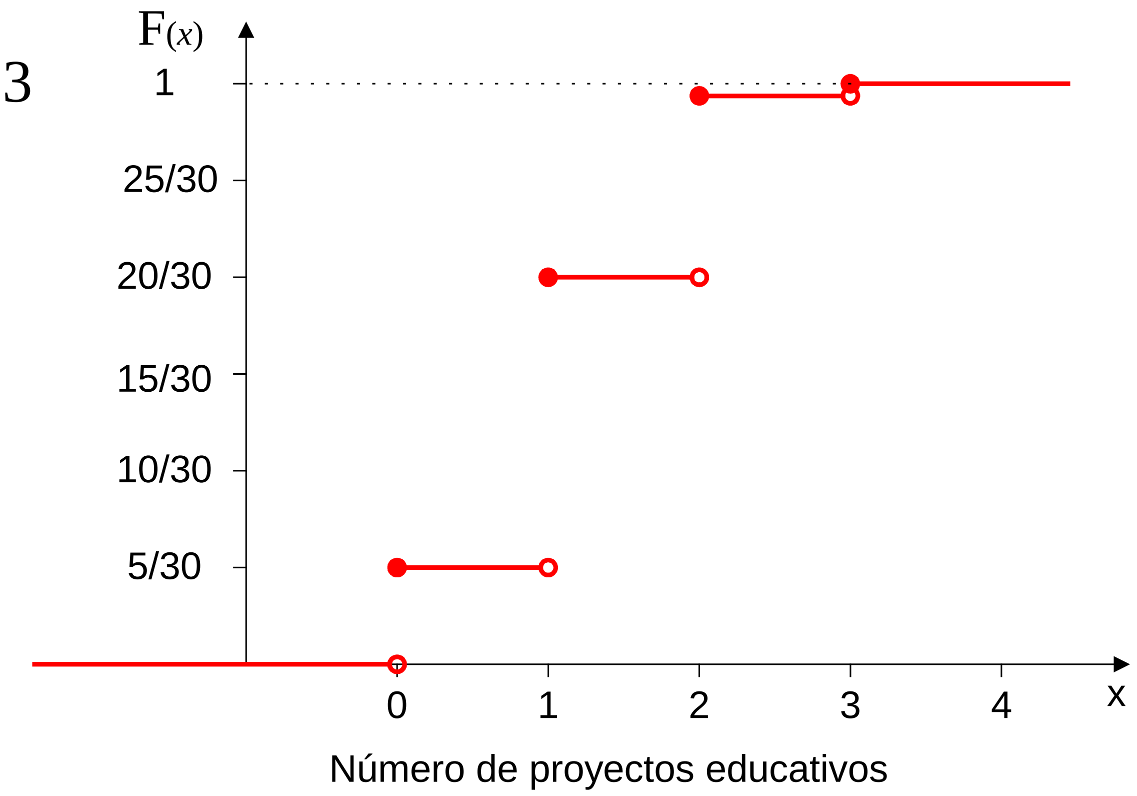
Para $x \geq 3$

$F(3)=P[\ X \leq 3\ ]= P[\ X < 3\ ]+P[\ X = \mathbf{3}\ ]= 29/30 + f(\mathbf{3}) = \mathbf{1}$

VARIABLES ALEATORIAS DISCRETAS. **Ejemplo**

$$F(x) = 0, \quad \text{si } x < 0$$
$$= 5/30, \quad \text{si } 0 \leq x < 1$$
$$= 20/30, \quad \text{si } 1 \leq x < 2$$
$$= 29/30, \quad \text{si } 2 \leq x < 3$$
$$= 1, \quad \text{si } x \geq 3$$

Gráfico de distribución acumulativa de probabilidades



VARIABLES ALEATORIAS CONTINUAS

Función de probabilidad

Una función *f(x)* es llamada función de probabilidad (o función de densidad) de la variable aleatoria continua X si cumple con las siguientes condiciones.

1. $f(x) \geq 0 \quad , \quad \forall\, x \in \mathbb{R}$

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

3. **Sea el evento A={ x / a $\leq$ $x \leq$ b}, luego,**

$$\mathrm{P}[\mathrm{A}] = \mathrm{P}[x \in \mathrm{A}] = \mathrm{P}[a \leq x \leq b] = \int_{a}^{b} f(x)dx$$

VARIABLES ALEATORIAS CONTINUAS. **Ejemplo.**

El tiempo de espera en una cola de servicios, en minutos, es una variable cuyo comportamiento aleatorio es definido por la siguiente función:

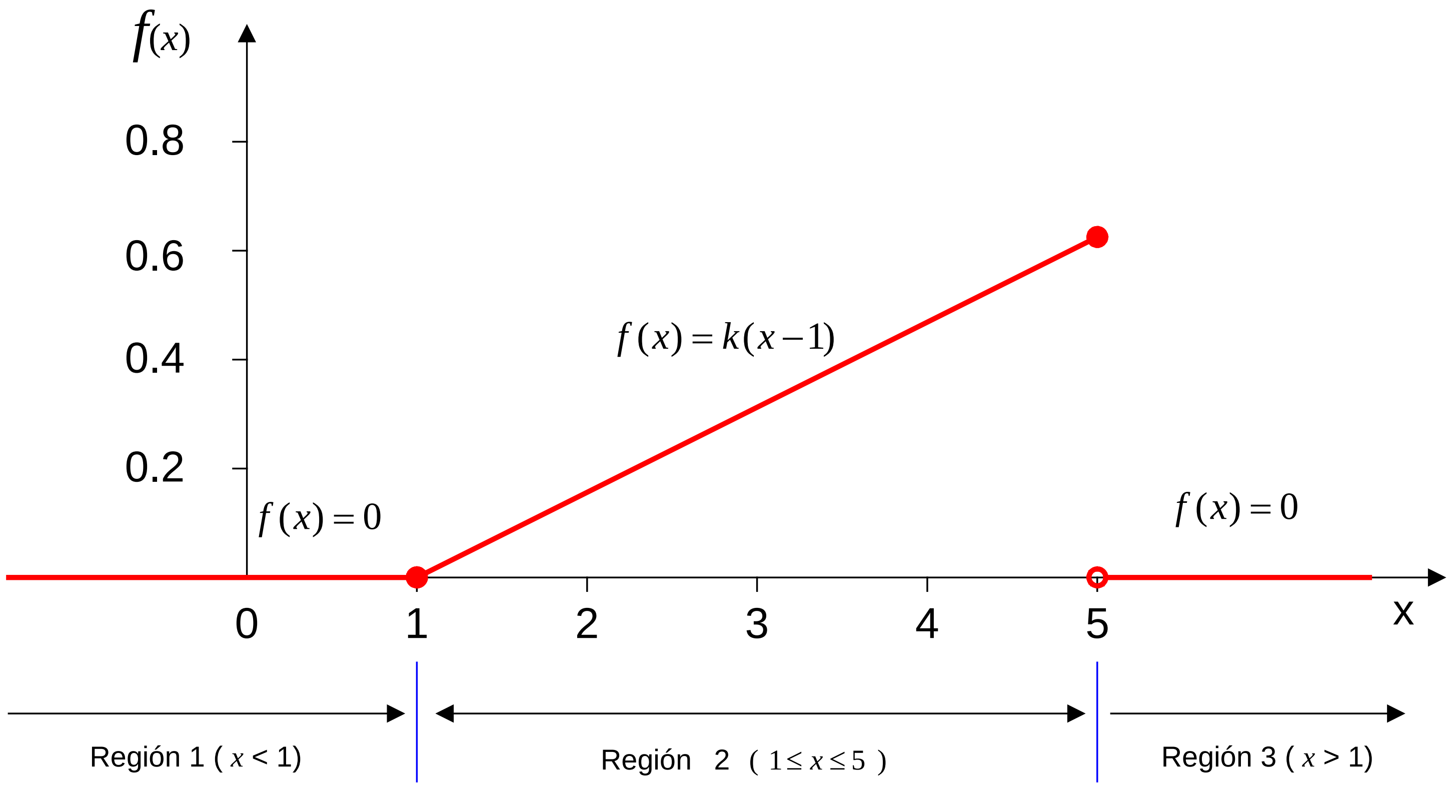
$$f(x) = k(x-1) \ , \quad si \quad 1 \leq x \leq 5$$

$$= 0 \qquad\qquad , \quad para \ \ otros \ \ valores \ \ de \ \ x$$

Halle el valor de "*k*" que hace que *f(x)* sea una función de densidad, y luego halle la probabilidad que un cliente elegido al azar espere más de tres minutos para ser atendido.

VARIABLES ALEATORIAS CONTINUAS. **Ejemplo.**

Distribución de los tiempos de espera



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

VARIABLES ALEATORIAS CONTINUAS. **Ejemplo.**

Analytics AoZ

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{x=1} f(x)dx + \int_{1}^{5} f(x)dx + \int_{x=5}^{\infty} f(x)dx$$

$$= 0 + \int_{1}^{5} k(x-1)dx + 0$$

$$= \int_{1}^{5} k(x-1)dx$$

$$= k\left.\frac{(x-1)^2}{2}\right|_{x=1}^{x=5} = 8k = 1 \quad \Rightarrow \quad k = \frac{1}{8}$$

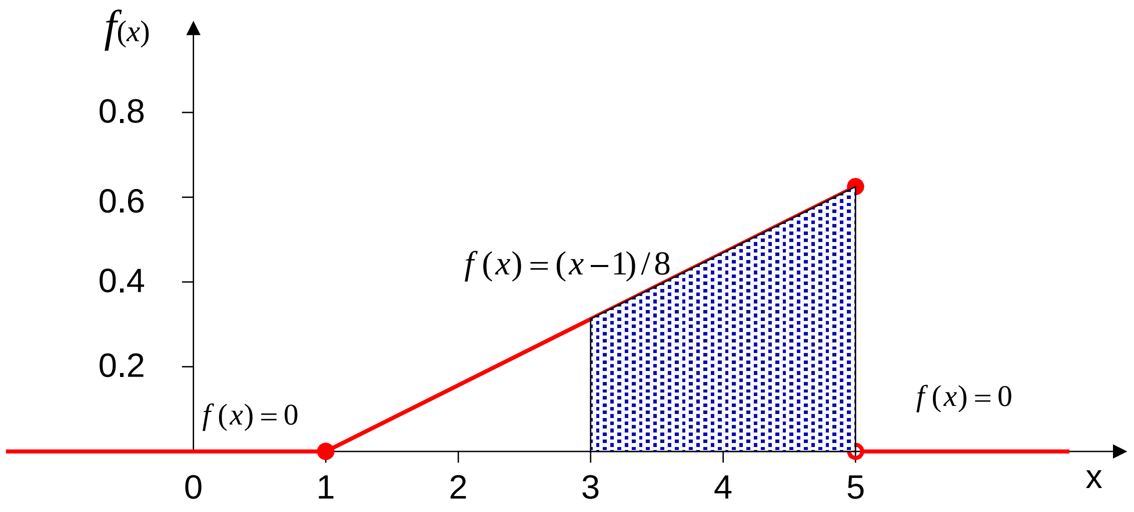
$$f(x) = (x-1)/8 \ , \quad si \quad 1 \le x \le 5$$

$$= 0 \qquad , \quad para \ \ otros \ \ valores \ \ de \ \ x$$

VARIABLES ALEATORIAS CONTINUAS. **Ejemplo.**

$$P\lfloor x > 3 \rfloor = ?$$

Distribución de los tiempos de espera



$$P[x > 3] = \int_{x=3}^{\infty} f(x)dx = \int_{x=3}^{x=5} f(x)dx + \int_{x=5}^{\infty} f(x)dx = \int_{x=3}^{5} \frac{x-1}{8}dx + 0$$

$$= \frac{(x-1)^2}{16}\Bigg|_{x=3}^{x=5} = \frac{1}{16}\left[(5-1)^2 - (5-3)^2\right] = 0.75$$

VARIABLES ALEATORIAS CONTINUAS

Analytics AoZ

Función de probabilidad acumulativa

Sea X una variable aleatoria continua que tiene como función de densidad *f(x)*; luego, la función de probabilidad acumulativa de la variable aleatoria X se define como:

$$F(x) = P[X \leq x] = \int_{-\infty}^{x} f(w)\,dw$$

Función de probabilidad acumulativa.

Propiedades

1. $F(-\infty) = 0$

2. $F(\infty) = 1$

3. $0 \leq F(x) \leq 1, \quad \forall \ x \in R$

4. F(x) es una función no decreciente.

5. $f(x) = \dfrac{d\ F(x)}{dx}$

VARIABLES ALEATORIAS CONTINUAS. **Ejemplo.**

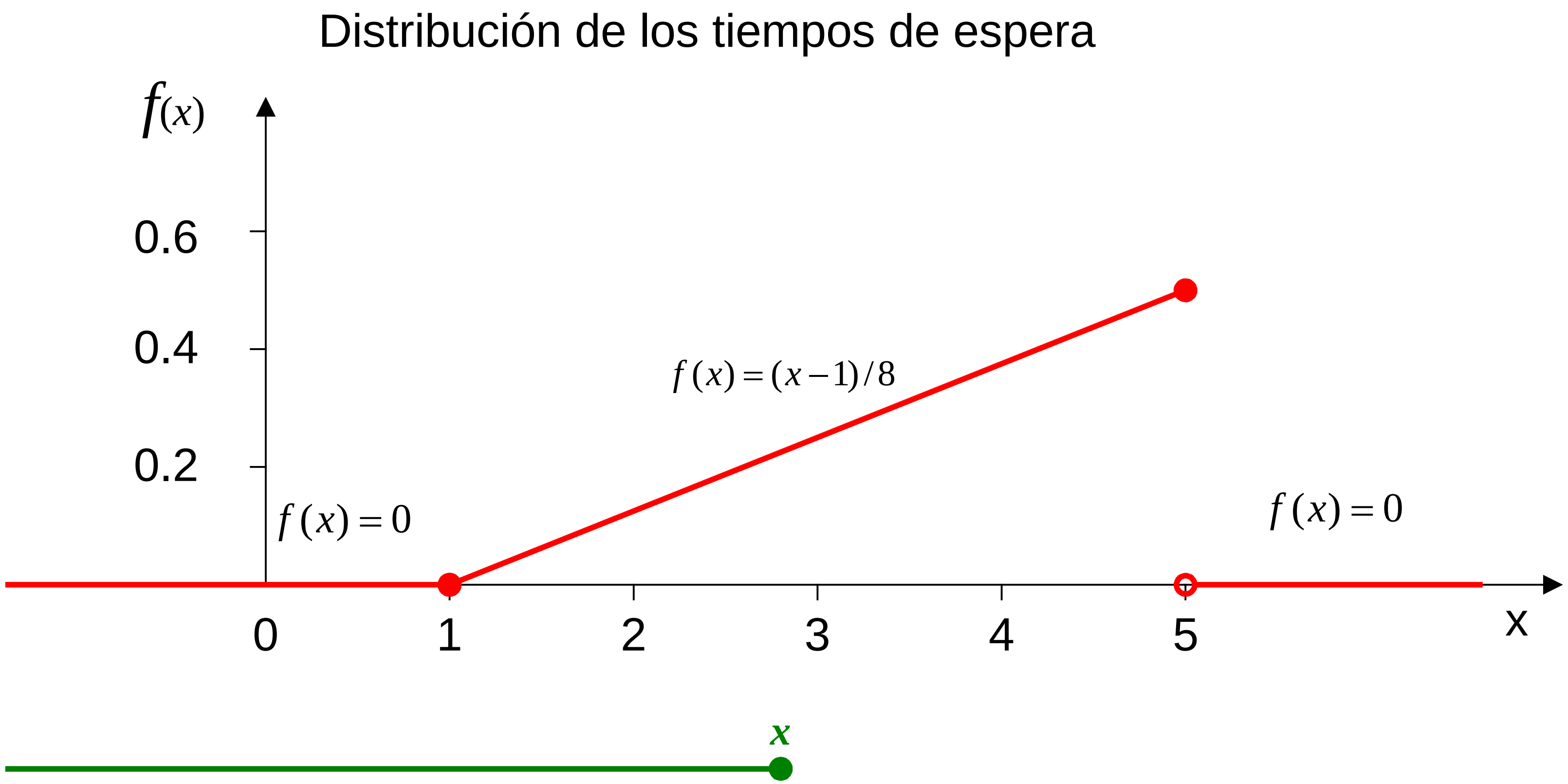
El tiempo de espera en una cola de servicios, en minutos, es una variable cuya función de densidad es:

$$f(x) = (x-1)/8 \ , \quad si \quad 1 \leq x \leq 5$$

$$= 0 \qquad \qquad , \quad para \ \ otros \ \ valores \ \ de \ \ x$$

Halle la función de distribución acumulativa de probabilidades de la variable. Luego halle la probabilidad que un cliente elegido al azar espere más de tres minutos para ser atendido.

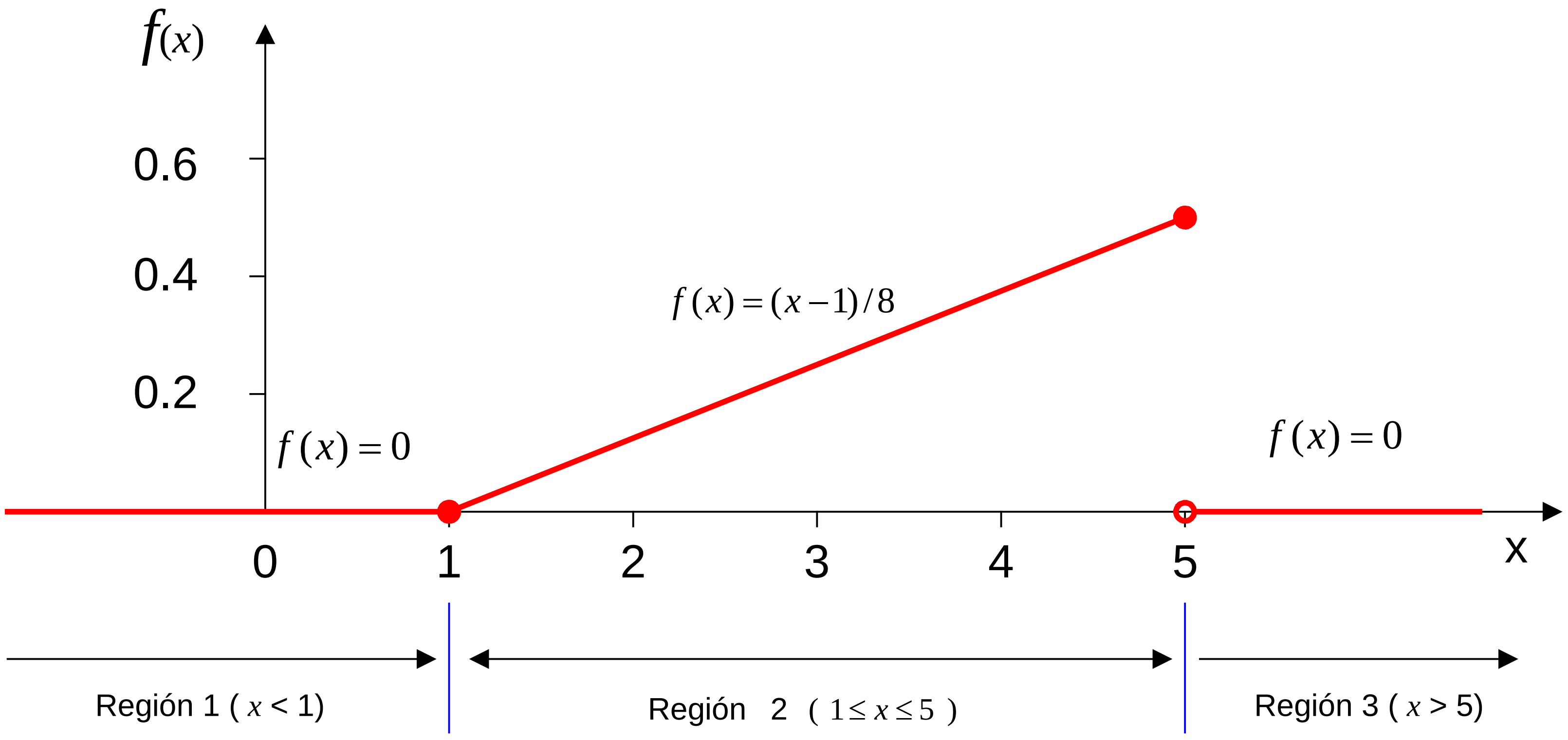
VARIABLES ALEATORIAS CONTINUAS. **Ejemplo.**



$$F(x) = P[X \leq x] = \int_{-\infty}^{x} f(w)dw$$

VARIABLES ALEATORIAS CONTINUAS. **Ejemplo.**

Distribución de los tiempos de espera

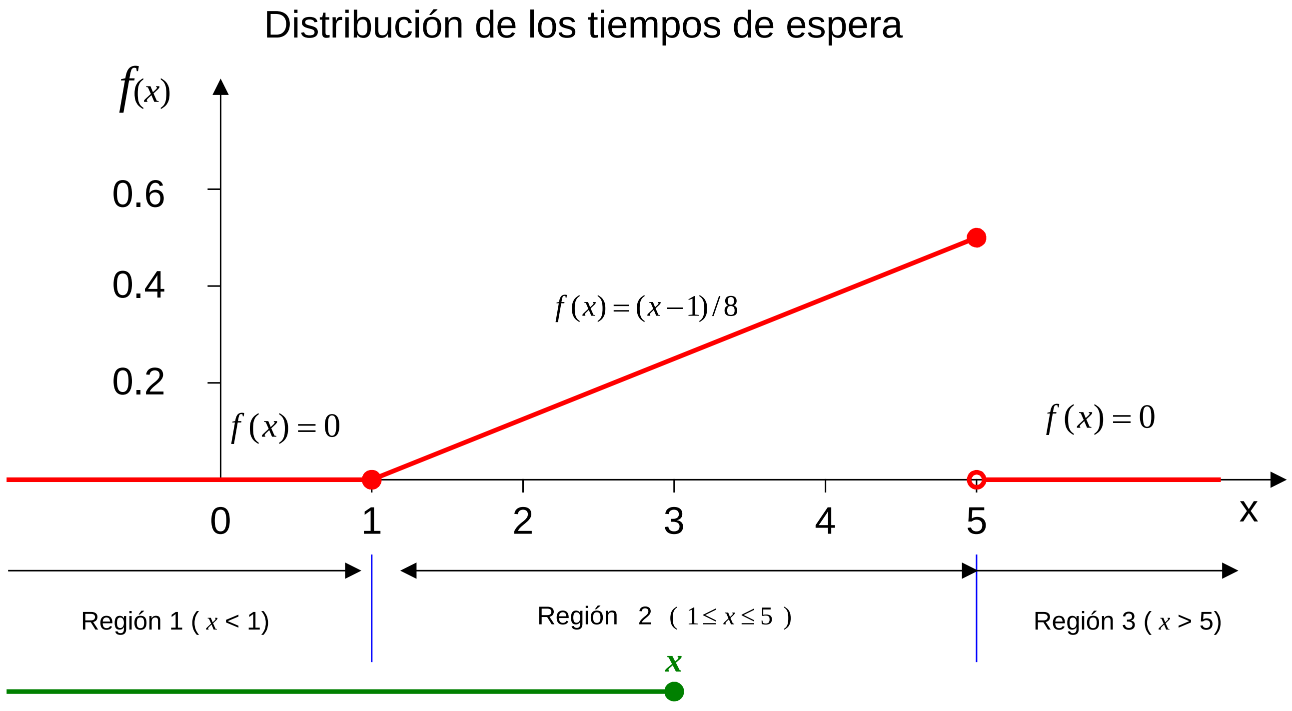


x

Para $x<1$

$$F(x)=P[X\le x]=\int_{-\infty}^{x} f(w)dw=\int_{-\infty}^{x}(0)dw=0$$

VARIABLES ALEATORIAS CONTINUAS. **Ejemplo.**

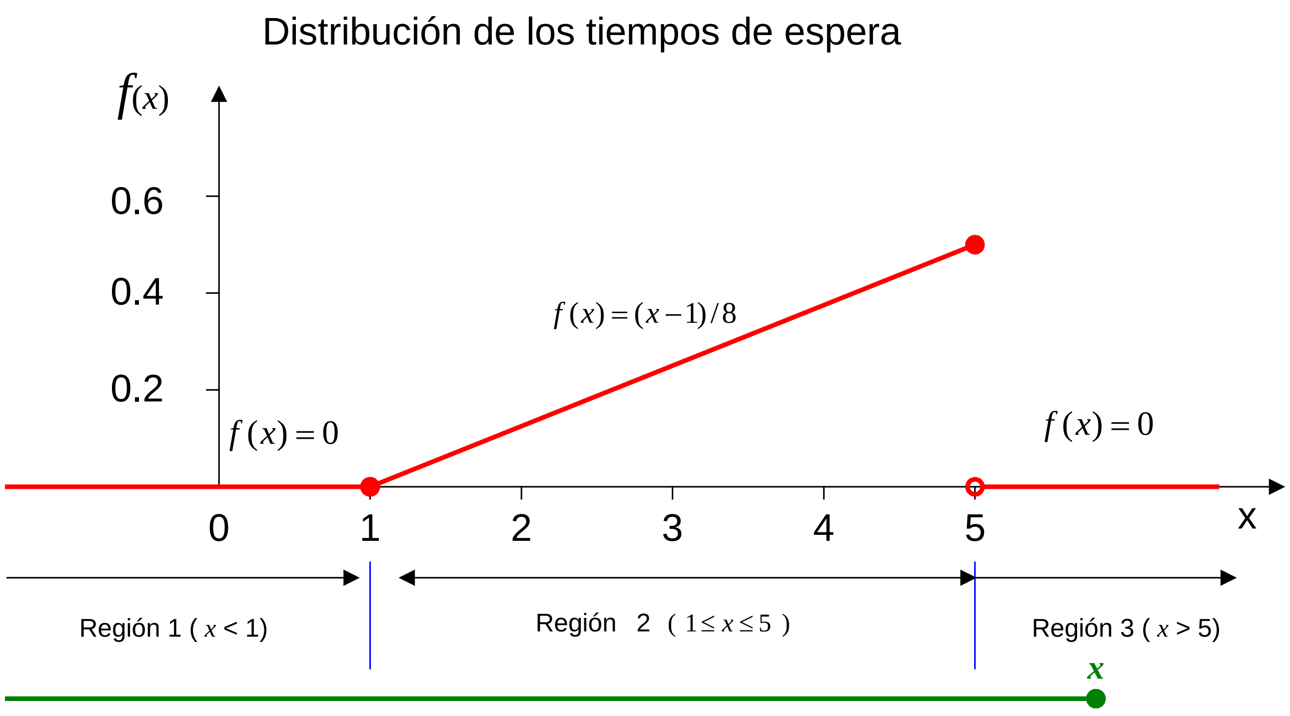


Para $1 \le x \le 5$

$$F(x)=P[X\le x]=\int_{-\infty}^{x} f(w)dw=\int_{-\infty}^{1} f(w)dw+\int_{w=1}^{x} f(w)dw$$

$$=0+\int_{1}^{x}\frac{w-1}{8}dw=\left.\frac{(w-1)^2}{16}\right|_{1}^{x}=\frac{(x-1)^2}{16}$$

VARIABLES ALEATORIAS CONTINUAS. **Ejemplo.**



Para $x > 5$

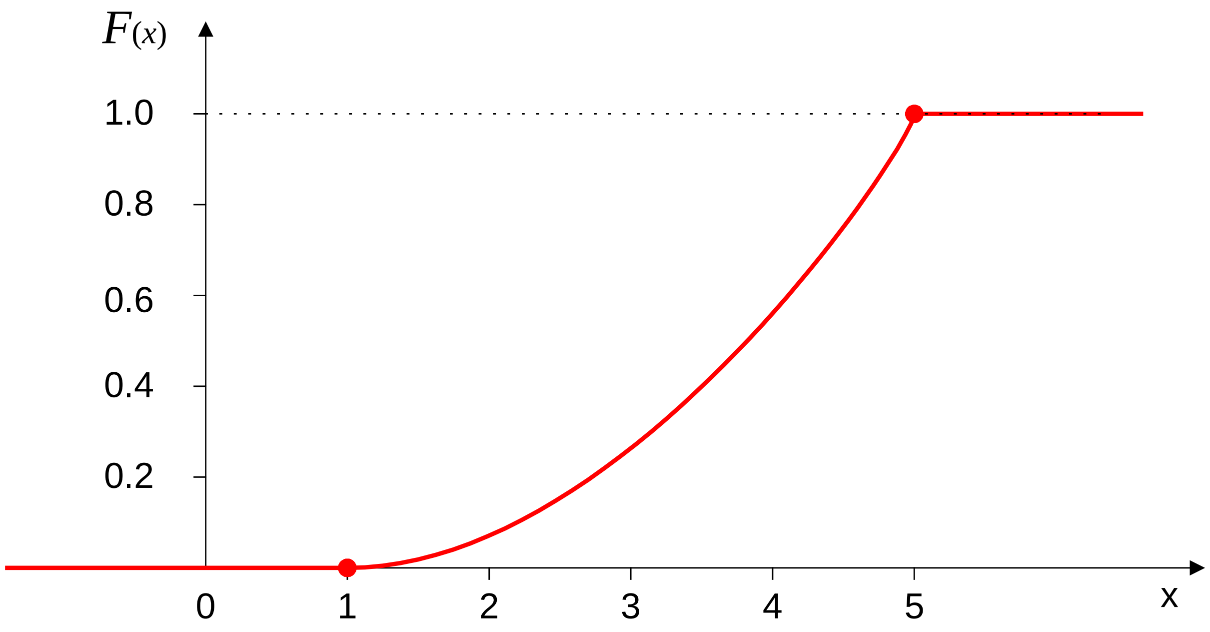
$$F(x) = P[X \leq x] = \int_{-\infty}^{x} f(w)dw = \int_{-\infty}^{1} f(w)dw + \int_{1}^{5} f(w)dw + \int_{w=5}^{x} f(w)dw$$

$$= 0 + 1 + 0 = 1$$

VARIABLES ALEATORIAS CONTINUAS. **Ejemplo.**

$$
\begin{aligned}
F(x) &= 0 && , \quad si \;\; x < 1 \\
&= \frac{(x-1)^2}{16} && , \quad si \;\; 1 \le x \le 5 \\
&= 1 && , \quad si \;\; x > 5
\end{aligned}
$$

Función de distribución acumulativa de probabilidades



VALOR ESPERADO

Valor esperado o esperanza matemática

Es el valor promedio que esperaríamos encontrar por la observación repetida de una variable aleatoria.

Si X es una variable aleatoria con función de probabilidad f(x), el valor esperado de la variable X se define como:

1. **Si X es una variable aleatoria discreta**

$$\mathrm{E}[\mathrm{X}] = \sum_{x_i \in R_X} x\, f(x_i)$$

2. **Si X es una variable aleatoria continua**

$$\mathrm{E}[\mathrm{X}] = \int_{-\infty}^{\infty} x\, f(x) dx$$

VALOR ESPERADO

Valor esperado o esperanza matemática de una función de una variable aleatoria

Sea X una variable aleatorias con función de probabilidad f(x), y sea Y=g(x) una función de la variable aleatoria X, luego,

1. **Si X es una variable aleatoria discreta**

$$\mathrm{E}\left[\mathrm{Y}\right] = \mathrm{E}\left[\, g(\mathrm{X}) \,\right] = \sum_{x_i \in R_X} g(x) \;\; f(x_i)$$

2. **Si X es una variable aleatoria continua**

$$\mathrm{E}\left[\mathrm{Y}\right] = \mathrm{E}\left[\, g(\mathrm{X}) \,\right] = \int_{-\infty}^{\infty} g(x) \;\; f(x)dx$$

VALOR ESPERADO

Media y variancia de una variable aleatoria

Sea X una variable aleatoria con función de probabilidad f(x) ; entonces, la media μ_X y la variancia de la variable aleatoria X se definen como:

$$\mu_X = E[X]$$

$$\sigma_X^2 = Var[X] = E[X - \mu_X]^2 = E[X^2] - \mu_X^2$$

VALOR ESPERADO

Propiedades de valores esperados

Si X, Y son dos variables aleatorias, y a, b son dos constantes:

1) E(a)=a
2) E(aX)= a E(X)
3) E(a X + b)= a E(X) + b
4) E(a X + b Y)= a E(X) + b E(Y)
5) Var(a) = 0
6) Var(a X)= a^2 Var(X)
7) Si X, Y son variables aleatorias independientes:

$$\mathrm{Var}[\mathrm{aX} \pm \mathrm{bY}] = a^2\ \mathrm{Var}[\mathrm{X}] + b^2\ \mathrm{Var}[\mathrm{Y}]$$

VALOR ESPERADO. **Ejemplo**

Si el número de proyectos educativos que son evaluados por semana es una variable aleatoria X que tiene como tabla de probabilidades:

X	0	1	2	3
$f(x)$	5/30	15/30	9/30	1/30

Halle la media y variancia de X.

$$\mathrm{E}[\mathrm{X}] = \sum x\, f(x)$$

$$= (0)(5/30) + (1)(15/30) + (2)(9/30) + (3)(1/30) = 1.2$$

$$\mathrm{E}[\mathrm{X}^2] = \sum x^2\, f(x)$$

$$= (0)^2(5/30) + (1)^2(15/30) + (2)^2(9/30) + (3)^2(1/30) = 2$$

$$\mu_X = \mathrm{E}[\mathrm{X}] = 1.2$$

$$\sigma_X^2 = \mathrm{E}[\mathrm{X}^2] - \mu_X^2 = 2 - (1.2)^2 = 0.56$$

VALOR ESPERADO. **Ejemplo**

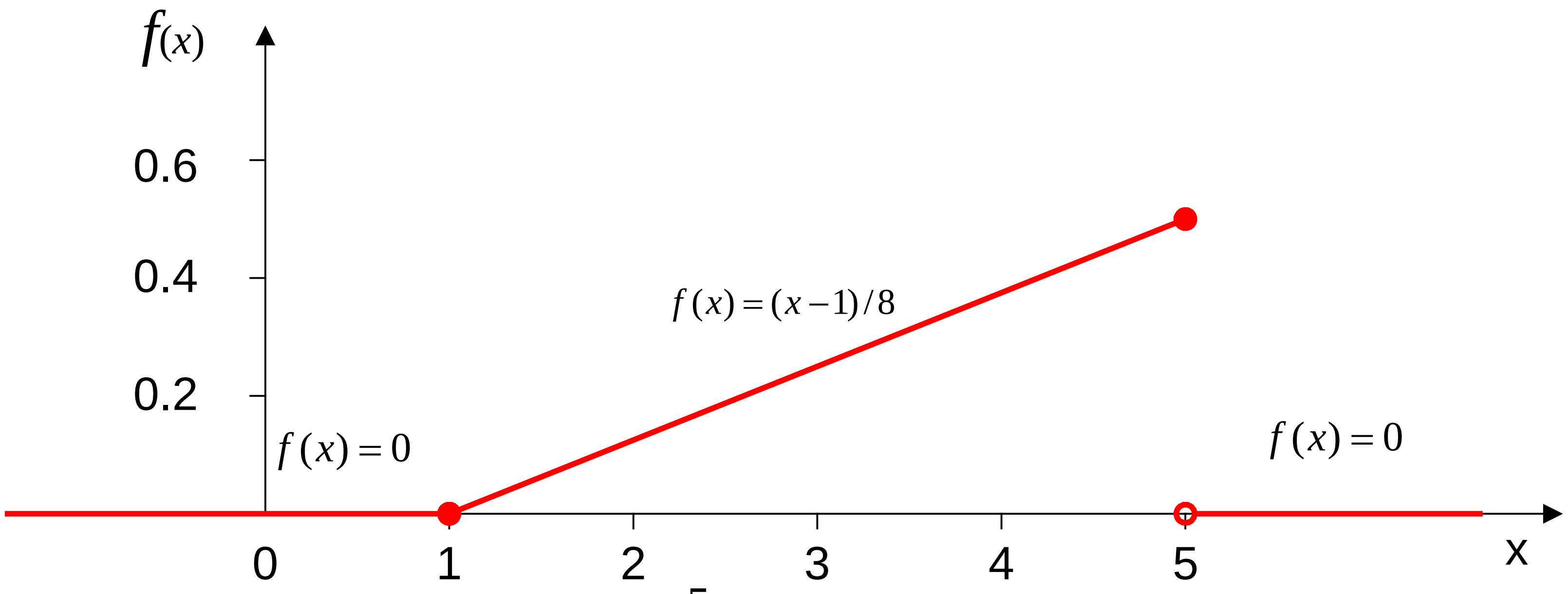
El tiempo de atención (en minutos), a los clientes de un establecimiento es una variable cuya función de densidad es:

$$f(x) = (x-1)/8 \ , \quad si \quad 1 \leq x \leq 5$$

$$= 0 \qquad\qquad , \quad para \ \ otros \ \ valores \ \ de \ \ x$$

Halle la media y variancia del tiempo de atención.

VALOR ESPERADO. **Ejemplo**



$$E[X] = \int_{-\infty}^{\infty} x\ f(x)dx = \int_{1}^{5} x\ \frac{x-1}{8} dx = 11/3$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2\ f(x)dx = \int_{1}^{5} x^2\ \frac{x-1}{8} dx = 43/3$$

$$\mu_X = E[X] = 11/3$$

$$\sigma_X^2 = E[X^2] - \mu_X^2 = \tfrac{43}{3} - (\tfrac{11}{3})^2 = \tfrac{8}{9}$$

Analytics AoZ

VALOR ESPERADO. **Ejemplo**

El tiempo de atención (X, en minutos), a los clientes de un establecimiento es una variable cuya función de densidad es:

$$f(x) = (x-1)/8 \ , \quad si \quad 1 \le x \le 5$$

$$= 0 \qquad\qquad , \quad para \ \ otros \ \ valores \ \ de \ \ x$$

Si el costo de atención (nuevos soles) es igual a : Y=5+1.2X, halle la media y variancia del costo de atención por cliente.

$$\mu_Y = E\left[5+1.2X\right] = 5+1.2\,E\left[X\right] = 5+1.2\,\mu_X$$

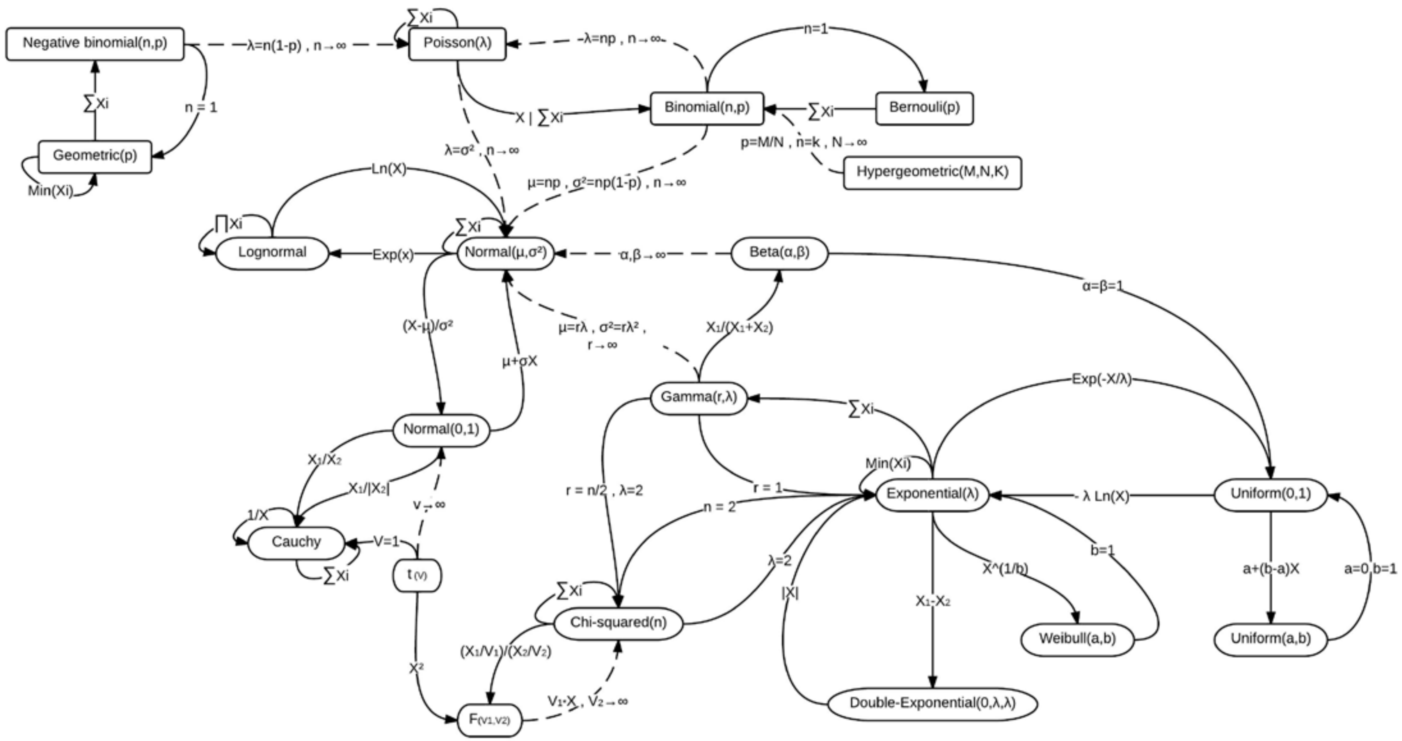
$$= 5+(1.2)(11/3) = 9.4 \qquad \text{nuevos soles}$$

$$\sigma_Y^2 = \mathrm{Var}\left[5+1.2X\right] = (1.2)^2\ \mathrm{Var}\left[X\right] = (1.2)^2\ \sigma_X^2$$

$$= (1.2)^2(8/9) = 1.28 \qquad (\text{nuevos soles})^2$$

DISTRIBUCIONES DE PROBABILIDAD y TRANSFORMACIONES

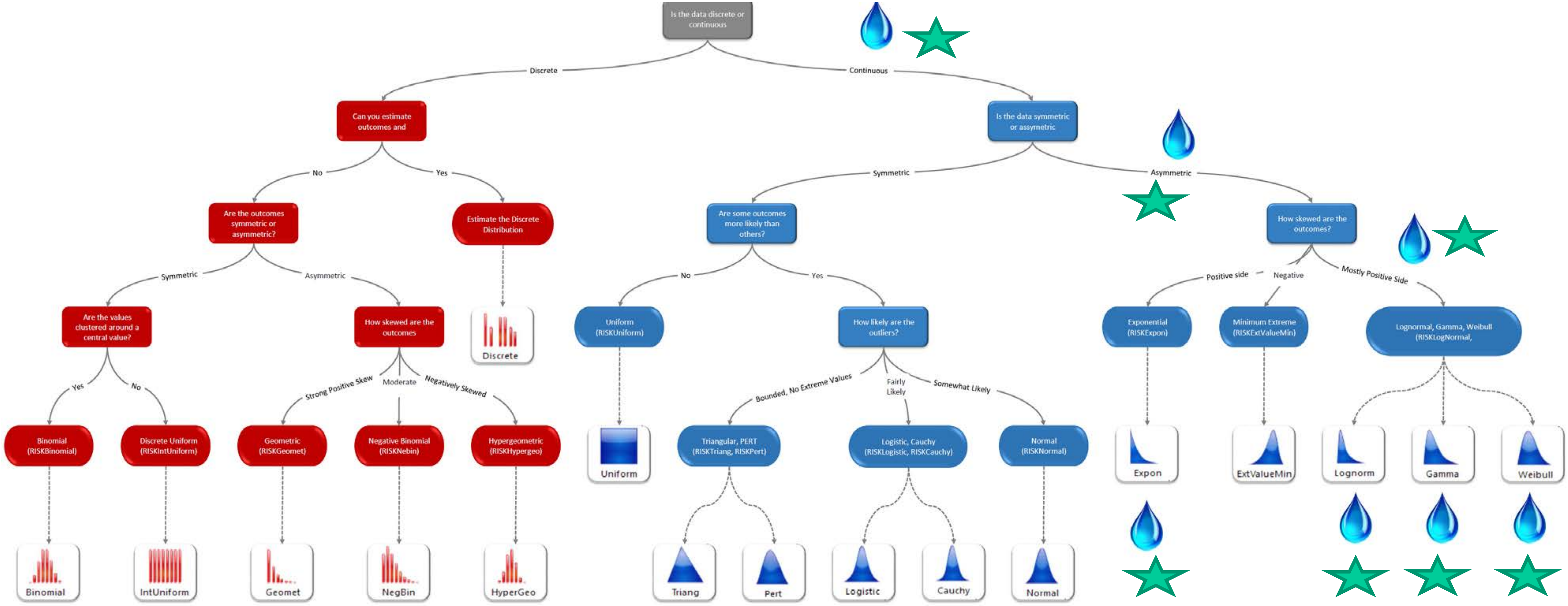
Distribution	Functions			
Beta	pbeta	qbeta	dbeta	rbeta
Binomial	pbinom	qbinom	dbinom	rbinom
Cauchy	pcauchy	qcauchy	dcauchy	rcauchy
Chi-Square	pchisq	qchisq	dchisq	rchisq
Exponential	pexp	qexp	dexp	rexp
F	pf	qf	df	rf
Gamma	pgamma	qgamma	dgamma	rgamma
Geometric	pgeom	qgeom	dgeom	rgeom
Hypergeometric	phyper	qhyper	dhyper	rhyper
Logistic	plogis	qlogis	dlogis	rlogis
Log Normal	plnorm	qlnorm	dlnorm	rlnorm
Negative Binomial	pnbinom	qnbinom	dnbinom	rnbinom
Normal	pnorm	qnorm	dnorm	rnorm
Poisson	ppois	qpois	dpois	rpois
Student t	pt	qt	dt	rt
Studentized Range	ptukey	qtukey	dtukey	rtukey
Uniform	punif	qunif	dunif	runif
Weibull	pweibull	qweibull	dweibull	rweibull
Wilcoxon Rank Sum Statistic	pwilcox	qwilcox	dwilcox	rwilcox
Wilcoxon Signed Rank Statistic	psignrank	qsignrank	dsignrank	rsignrank



Fuente: http://www.stat.umn.edu/geyer/old/5101/rlook.html

DISTRIBUCIONES DE PROBABILIDAD

Analytics AoZ



Fuente: https://www.palisade.com/

DISTRIBUCIONES DE PROBABILIDAD

Distribuciones discretas

Son las funciones de probabilidad asociadas a variables aleatorias discretas, que se generan mediante procesos de conteo sobre las veces que se repite un suceso.

Cuando se elige al azar un elemento, se averigua si cumple o no con cierta condición, para luego contar el número de elementos que si cumplen con la condición en análisis.

Distribuciones discretas. Pruebas de Bernoulli

Un experimento aleatorio es llamado una prueba o ensayo de Bernoulli si cumple las siguientes condiciones:

1. Para cada prueba o ensayo se define un espacio muestral con solo dos resultados posibles: Exito (E) y Fracaso (F), donde:

$$P[E]=\pi \quad y \quad P[F]= 1 - P[E]= 1 - \pi$$

2. La probabilidad de éxito (π) se mantiene constante de prueba a prueba.

3. Las pruebas se consideran que son independientes.

Distribuciones discretas

Distribución Bernoulli o binomial puntual

Analytics AoZ

Una variable aleatoria discreta X tiene una distribución Bernoulli o binomial puntual si su función de probabilidad es dada por:

$$f(x) = \pi^{x}(1-\pi)^{1-x} \quad , \text{ si } \quad x = 0,1$$

$$= 0 \qquad\qquad , \text{ de otro modo}$$

Donde: π = Probabilidad de éxito

X = Número de éxitos en una prueba de Bernoulli

Además, la media y variancia de la variable aleatoria X son:

$$\mu_{X} = \mathrm{E}[X] = \pi$$

$$\sigma_{X}^{2} = \mathrm{Var}[X] = \mathrm{E}[X\text{-}\mu_{X}]^{2} = \pi(1-\pi)$$

Distribuciones discretas

Distribución Binomial

Una variable aleatoria discreta X tiene una distribución Binomial si su función de probabilidad es dada por:

$$f(x) = \binom{n}{x} \pi^{x} (1-\pi)^{n-x} \quad , \quad si \quad x = 0,1,\cdots,n$$

$$= 0 \qquad\qquad , \quad de\ otro\ modo$$

Donde:

π = Probabilidad de éxito

n = Número de pruebas de Bernoulli o
tamaño de una muestra con reemplazo

X = Número de éxitos en "n" pruebas de Bernoulli

Además, la media y variancia de la variable aleatoria X son:

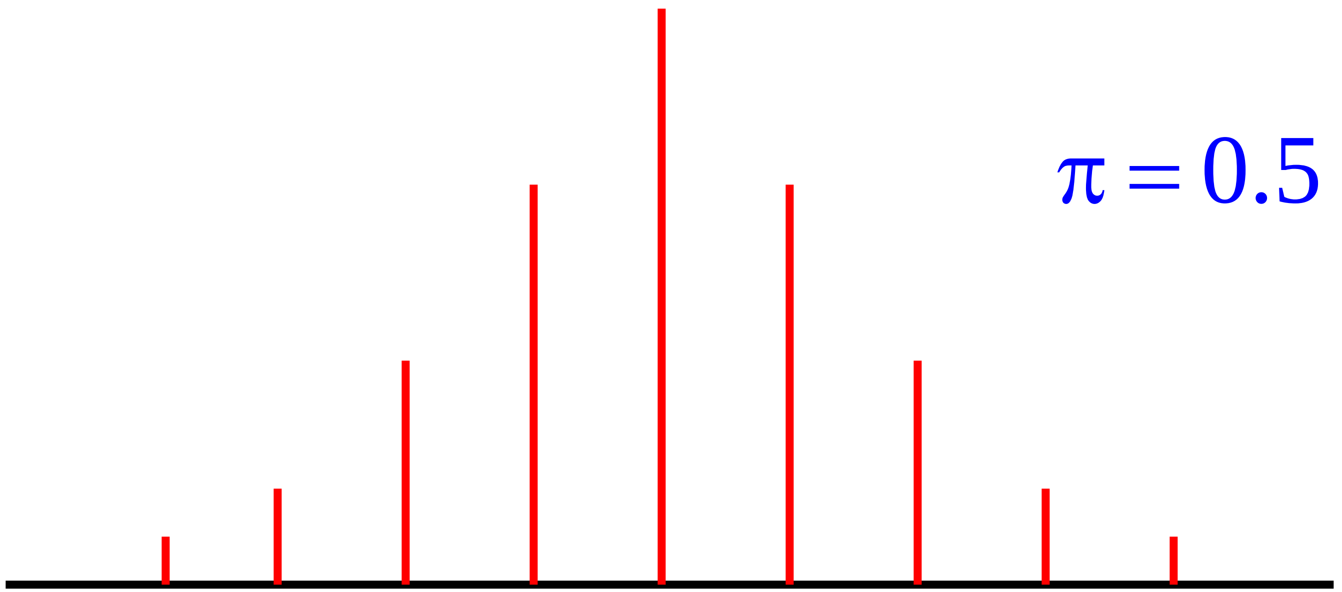
$$\mu_X = \mathrm{E}[X] = n\,\pi$$

$$\sigma_X^2 = \mathrm{Var}[X] = \mathrm{E}[X - \mu_X]^2 = n\,\pi\,(1-\pi)$$

Analytics AoZ

Distribuciones discretas

Distribución Binomial



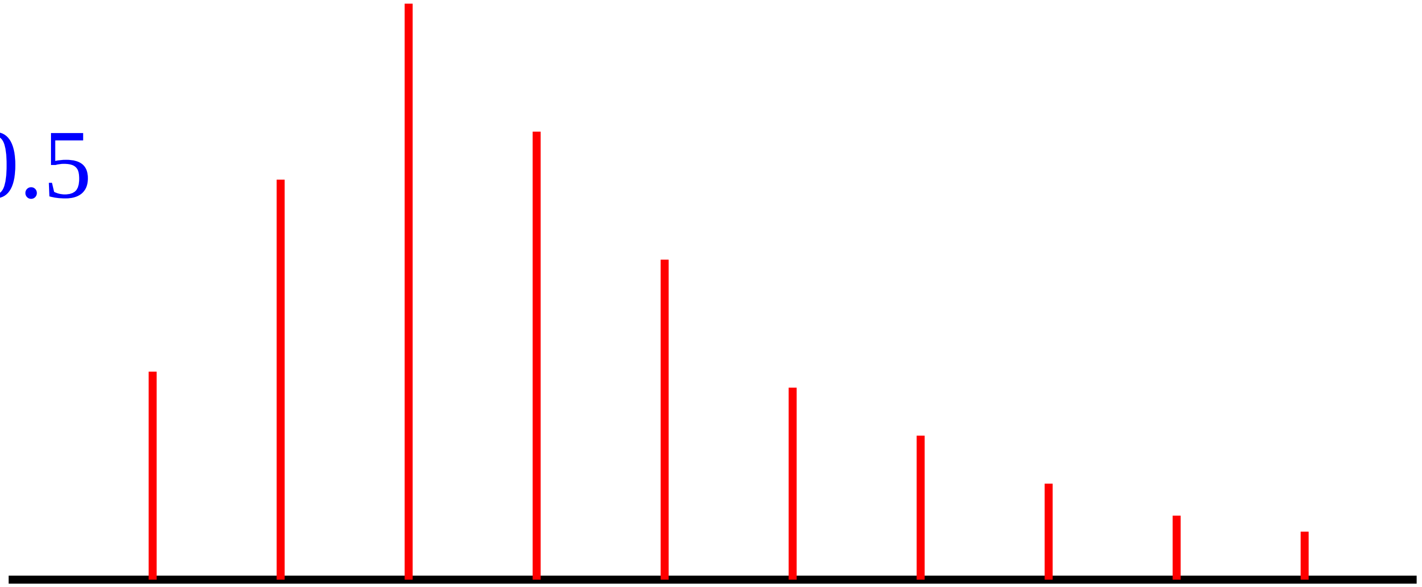
$$f(x) = \binom{n}{x} \pi^{x} (1-\pi)^{n-x} \quad , \quad si \quad x = 0, 1, \cdots, n$$

$$= 0 \quad , \quad de\ otro\ modo$$

Distribuciones discretas

Distribución Binomial

$$\pi < 0.5$$

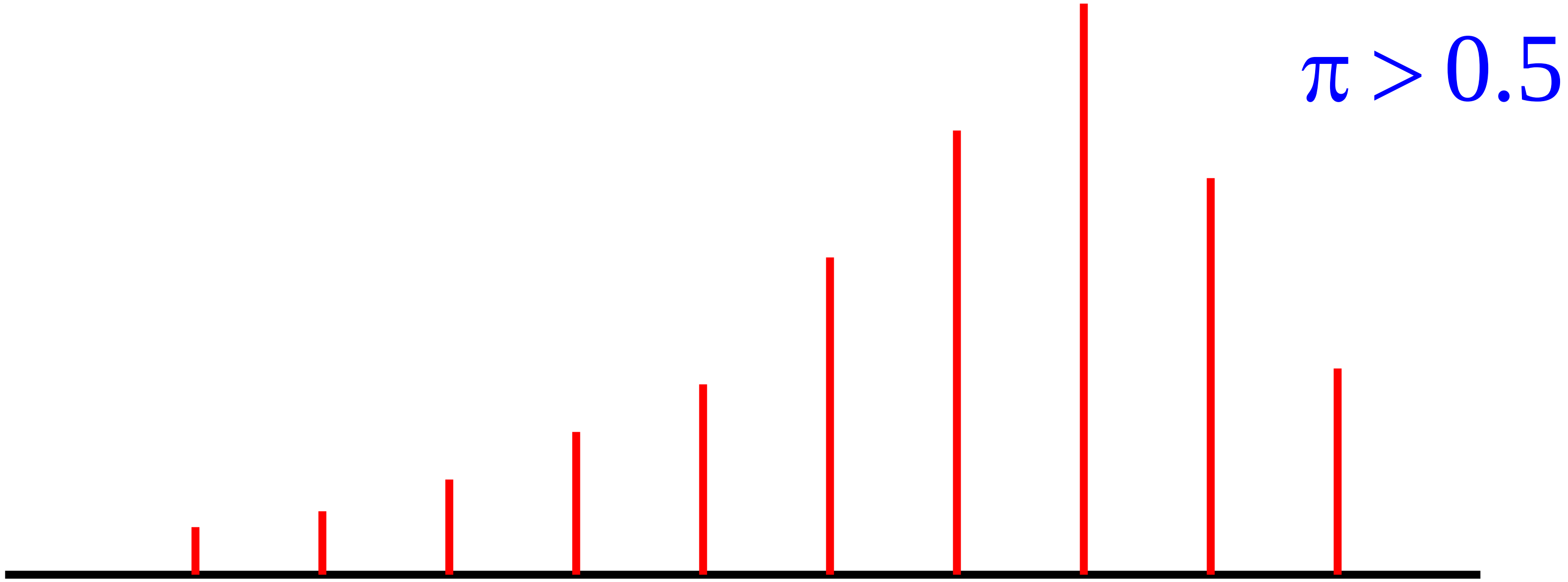


$$f(x) = \binom{n}{x} \pi^{x} (1-\pi)^{n-x} \quad , \quad si \quad x = 0, 1, \cdots, n$$

$$= 0 \quad , \quad de\ otro\ modo$$

Distribuciones discretas

Distribución Binomial



$$f(x) = \binom{n}{x} \pi^x (1-\pi)^{n-x} \quad , \quad si \quad x = 0, 1, \cdots, n$$

$$= 0 \quad , \quad de\ otro\ modo$$

Distribuciones discretas

Distribución Binomial. Ejemplo

Analytics AoZ

EL 10% de fragmentos de roca obtenidos por una taladro son defectuosos (polvo). Si elige una muestra aleatoria con reemplazo de 6 artículos y se define la variable X como el número de artículos defectuosos elegidos,

a) Determinar la probabilidad que al menos un fragmento sea defectuoso

$$\pi = P\lfloor Exito \rfloor = P\lfloor Defectuso \rfloor = 0.1$$

$$f(x) = \binom{6}{x}(0.1)^x(1-0.1)^{6-x} \quad , \quad si \quad x = 0,1,\cdots,6$$

$$= 0 \qquad\qquad , \quad de\ otro\ modo$$

$$P\lfloor X \geq 1 \rfloor = 1 - P\lfloor X < 1 \rfloor = 1 - P\lfloor X = 0 \rfloor = 1 - f(o)$$

$$= 1 - \binom{6}{0}(0.1)^0(1-0.1)^{6-0} = 0.468559$$

Analytics AoZ

b) Halle el valor del coeficiente de variabilidad de X

$$\mu_X = n\pi = (6)(0.1) = 0.6$$

$$\sigma_X^2 = n\pi(1-\pi) = (6)(0.1)(1-0.1) = 0.54$$

$$CV_X = \left(\frac{\sigma_X}{\mu_X}\right)(100) = \left(\frac{\sqrt{0.54}}{0.6}\right)(100) = 122.474487$$

Distribuciones discretas

Distribución de Poisson

Analytics AoZ

Una variable aleatoria discreta X tiene una distribución de Poisson si su función de probabilidad es dada por:

$$f(x) = \frac{e^{-\mu}\,\mu^{x}}{x!} \quad , \quad si \quad x = 0,1,2,\cdots$$

$$= 0 \quad , \quad de\ otro\ modo$$

Donde:

X = Número de éxitos obtenidos en un período o unidad de evaluación

$\mu = \lambda$ t

λ = Razón media de ocurrencia por periodo

t = Periodo o unidad de evaluación

Además, la media y variancia de la variable aleatoria X son:

$$\mu_X = \mathrm{E}\left[X\right] = \mu$$

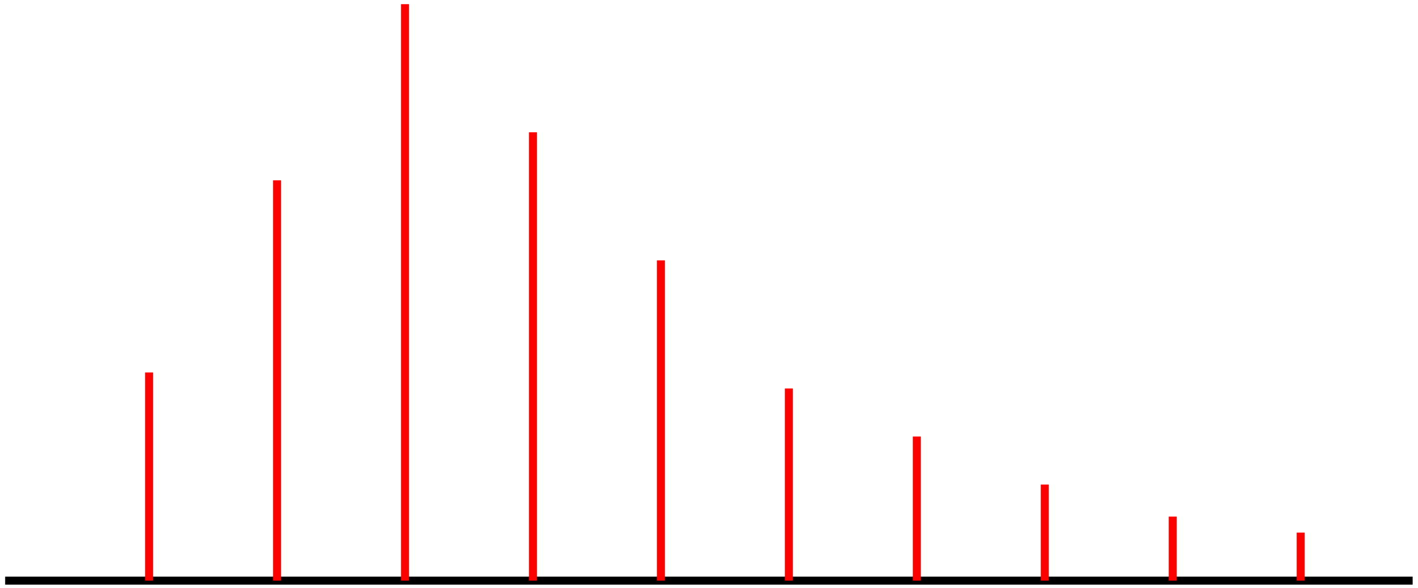
$$\sigma_X^2 = \mathrm{Var}\left[X\right] = \mathrm{E}\left[X - \mu_X\right]^2 = \mu$$

Para la aplicación de esta distribución es necesario tener en cuenta los siguientes supuestos:

1. La probabilidad de éxito se mantiene constante y es pequeña.

2. El número de éxitos en cualquier intervalo, periodo o unidad de evaluación es independiente del número de éxitos en cualquier otro intervalo, periodo o unidad de evaluación.

Distribuciones discretas. Distribución de Poisson



$$f(x) = \frac{e^{-\mu}\,\mu^{x}}{x!} \quad , \quad si \quad x = 0,1,2,\cdots$$

$$= 0 \quad , \quad de\ otro\ modo$$

Distribuciones discretas

Distribución de Poisson Ejemplo

Los perforaciones de un taladro se hacen a razón de 4 metros por cada 5 minutos. Si se elige al azar un intervalo de 2 minutos, hallar la probabilidad que se taladro al menos dos metros de perforación.

$$\lambda = \frac{4}{5}\frac{taladros}{\text{minuto}} = 0.8 \quad {}^{ta}/_{m}$$

$$t = 2 \ \text{minutos}$$

$$\mu = \lambda\, t = (0.8)(2) = 1.6 \quad taladros$$

$$f(x) = \frac{e^{-1.6}\,(1.6)^x}{x!} \quad , \quad si \quad x = 0,1,2,\cdots$$

$$= 0 \qquad , \quad de\ otro\ modo$$

$$P[x \geq 2] = 1 - P[x < 2] = 1 - f(0) - f(1) = 1 - \frac{e^{-1.6}\,(1.6)^0}{0!} - \frac{e^{-1.6}\,(1.6)^1}{1!}$$

$$= 0.475069$$

Aproximación de Poisson a una Distribución Binomial

La distribución de Poisson se puede utilizar para aproximar las probabilidades de una distribución binomial cuando n tiende a infinito ($n \to \infty$) y π tiende a cero ($\pi \to 0$); es decir, cuando $\boldsymbol{n\,\pi < 5}$. Para esto se debe considerar: $\mu = n\,\pi$ ($\lambda = \pi$, $t = n$).

Distribuciones discretas

Distribución de Poisson Ejemplo

Suponga que el 5% de las rocas de la Fm. Ananea contienen oro. Si se eligen al azar y con reemplazo 90 rocas, hallar la probabilidad que en al menos 3 rocas se encuentren con oro.

Si se define Éxito = { Una roca contienen oro }, se tiene: π=P[E]=0.05 y 1-π=P[F]=0.95. Luego, la función de probabilidad de X (número de rocas con oro) será:

$$f(x) = \binom{90}{x}(0.05)^{x}(1-0.05)^{90-x} \quad , \quad si \quad x = 0,1,\cdots,90$$

$$= 0 \qquad\qquad , \quad de\ otro\ modo$$

$$f(x) = \binom{90}{x}(0.05)^x(1-0.05)^{90-x} \quad , \quad si \quad x = 0,1,\cdots,90$$
$$= 0 \quad , \quad de\ otro\ modo$$

Ahora, como μ= n π=(90)(0.05)=4.5<5; luego, las probabilidades de X se pueden aproximar mediante una distribución de Poisson con parámetro μ=4.5. Es decir,

$$f(x) \approx \frac{e^{-4.5}\, 4.5^x}{x!} \quad , \quad si \quad x = 0,1,2,\cdots$$
$$= 0 \quad , \quad de\ otro\ modo$$

$$\mathrm{P}\lfloor X \geq 3 \rfloor = 1 - \mathrm{P}\lfloor X < 3 \rfloor \approx 1 - f(0) - f(1) - f(2)$$
$$\approx 1 - \frac{e^{-4.5}(4.5)^0}{0!} - \frac{e^{-4.5}(4.5)^1}{1!} - \frac{e^{-4.5}(4.5)^2}{2!}$$

Distribuciones discretas.
Distribución Geométrica

Analytics AoZ

Una variable aleatoria discreta X tiene una distribución geométrica si su función de probabilidad es dada por:

$$f(x) = \pi(1-\pi)^{x-1} \quad , \quad si \quad x = 1, 2, \cdots$$
$$= 0 \quad , \quad de\ otro\ modo$$

Donde:

X = Número de pruebas de Bernoulli hasta obtener el primer éxito

π = Probabilidad de éxito en una prueba Bernoulli

Además, la media y variancia de la variable aleatoria X son:

$$\mu_X = \mathrm{E}[X] = 1/\pi$$

$$\sigma_X^2 = \mathrm{Var}[X] = \mathrm{E}[X - \mu_X]^2 = (1-\pi)/\pi^2$$

Distribuciones discretas

Distribución Geométrica Ejemplo

Suponga que el mejor geólogo del mundo tiene una probabilidad 0.04 de no encontrar un yacimiento (fallar), y que cuando ello ocurre es necesario reemplazarlo por uno nuevo. Determinar la media y el coeficiente de variabilidad del número de veces que puede ser utilizado el geólogo para descubrir yacimientos, además del coeficiente de variación.

Si se define Éxito={El geólogo falla al encontrar un yacimiento}, se tiene: π=P[E]=0.04 y 1-π=P[F]=0.96. Luego, la función de probabilidad de X (número de veces que puede ser utilizado el geólogo para encontrar un yacimiento) será:

$$\begin{aligned} \boldsymbol{f}(\boldsymbol{x}) &= (0.04)(0.96)^{x-1} \quad , \quad \boldsymbol{si} \quad \boldsymbol{x} = 1,2,\cdots \\ &= 0 \quad , \quad \textit{de otro modo} \end{aligned}$$

$$\mu_X = \mathrm{E}[X] = 1/0.04 = 25$$

$$\sigma_X^2 = \mathrm{Var}[X] = \mathrm{E}[X-\mu_X]^2 = (1-\pi)/\pi^2 = 0.96/(0.04)^2 = 600$$

$$CV_X = \frac{\sqrt{600}}{25}(100) = 97.9796\%$$

Distribuciones discretas.
Distribución Hipergeométrica

Analytics AoZ

Sea una población de tamaño N, donde hay A elementos que tienen una característica W definida como éxito, y B elementos que no tienen la característica W, siendo N=A+B. Si de dicha población se toma una muestra aleatoria de tamaño n y se define la variable aleatoria X como el número de éxitos obtenidos; entonces, la variable X tendrá como función de probabilidad:

$$f(x) = \frac{\binom{A}{x}\binom{B}{n-x}}{\binom{N}{n}} \quad , \quad si \quad x = 0,1,2,\cdots,n$$

$$= 0 \qquad\qquad , \quad de\ otro\ modo$$

Donde: X = Número de éxitos obtenidos en una muestra sin reemplazo de tamaño n.

Analytics AoZ

Además, la media y variancia de la variable aleatoria X son:

$$\mu_X = E[X] = \frac{n\,A}{N}$$

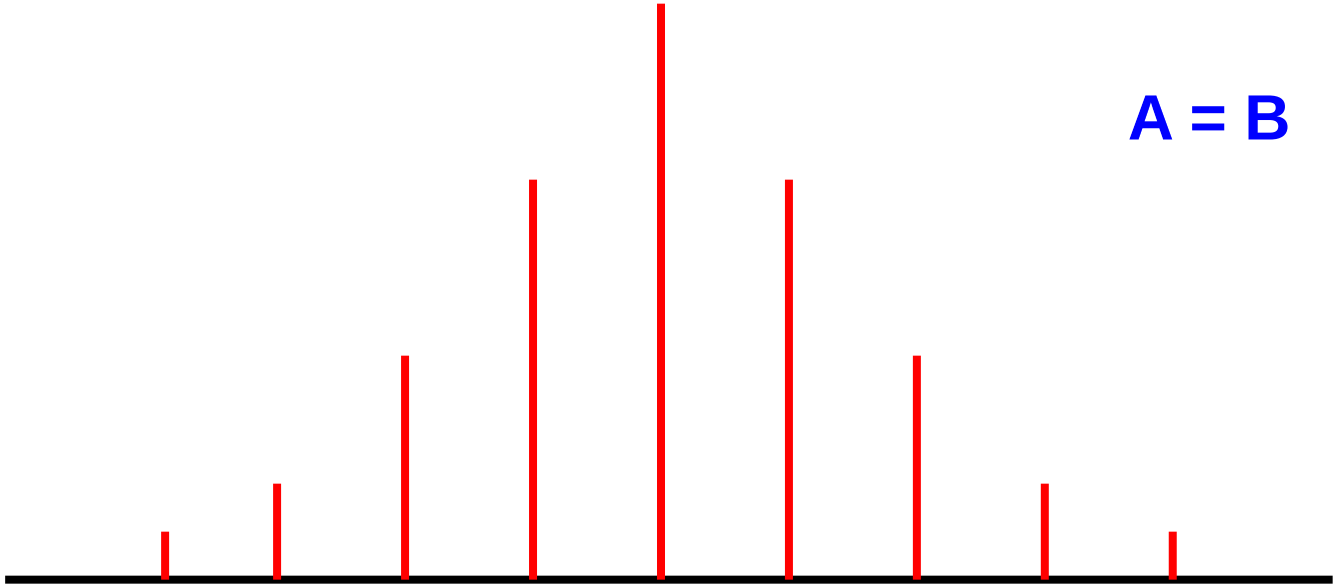
$$\sigma_X^2 = Var[X] = \frac{n\,A\,B}{N^2}\left[\frac{N-n}{N-1}\right]$$

Nota

Si el número de elementos de la población (N) es grande y la fracción de muestreo (f=n/N) es pequeña; es decir, si f<0.05, entonces las probabilidades de la distribución hipergeométrica se pueden aproximar mediante la distribución binomial, para lo cual se considera: π= A/N.

Distribuciones discretas

Distribución Hipergeométrica



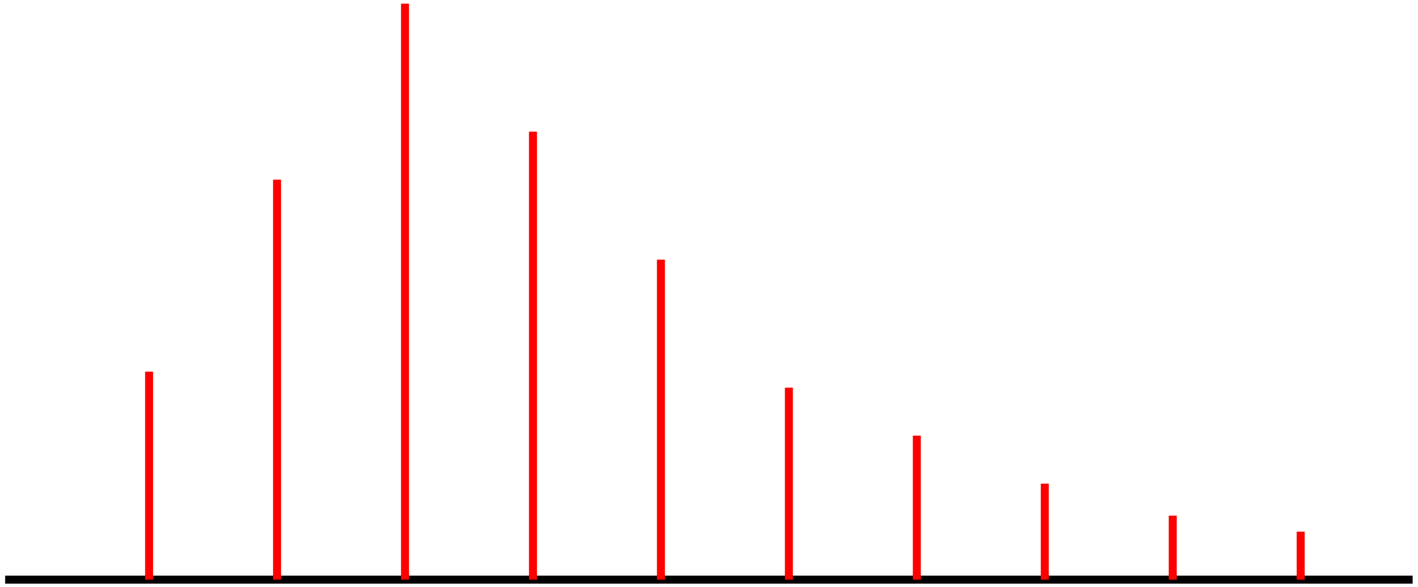
$$f(x) = \frac{\binom{A}{x}\binom{B}{n-x}}{\binom{N}{n}} \quad , \quad si \quad x = 0,1,2,\cdots,n$$

$$= 0 \qquad\qquad , \quad de\ otro\ modo$$

Distribuciones discretas

Distribución Hipergeométrica

A < B

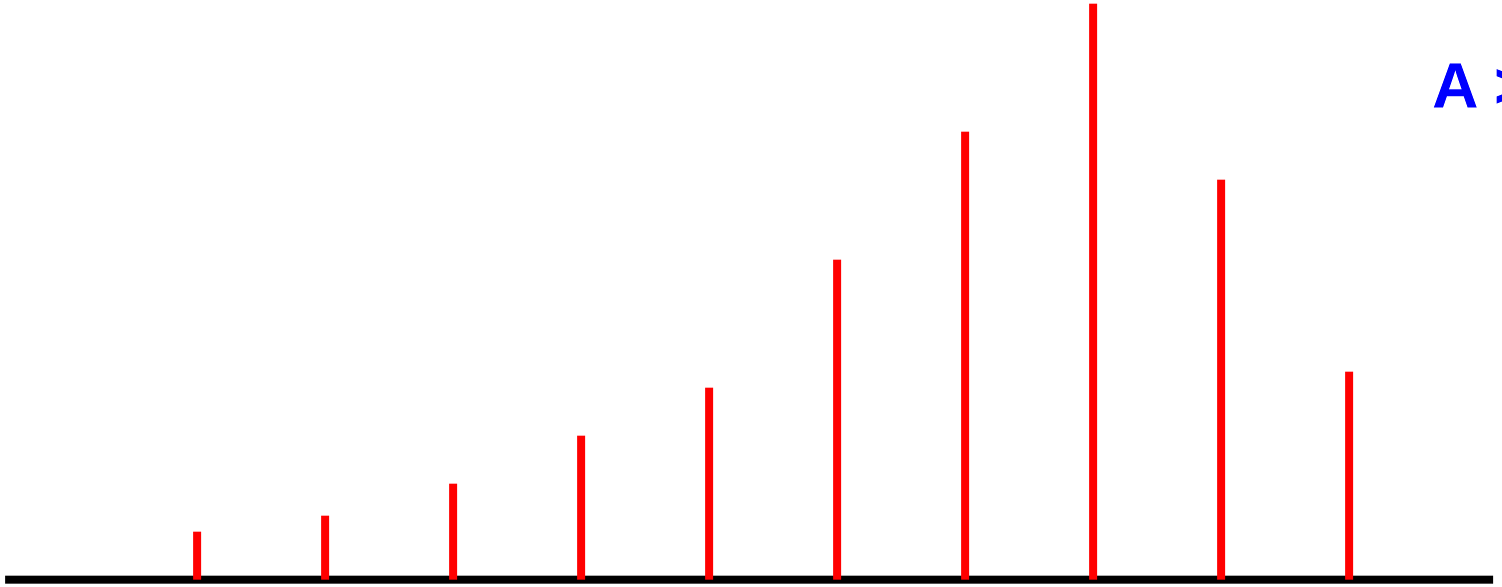


$$f(x)=\frac{\binom{A}{x}\binom{B}{n-x}}{\binom{N}{n}}\quad,\quad si\quad x=0,1,2,\cdots,n$$

$$=0\qquad\qquad,\quad de\ otro\ modo$$

Distribuciones discretas

Distribución Hipergeométrica



A > B

$$f(x) = \frac{\binom{A}{x}\binom{B}{n-x}}{\binom{N}{n}} \quad , \quad si \quad x = 0,1,2,\cdots,n$$

$$= 0 \qquad\qquad , \quad de\ otro\ modo$$

Distribuciones discretas.

Distribución Hipergeométrica   Ejemplo

Suponga que en un proceso de control de calidad de base de datos geológico se inspecciona un 10 tablas de base de datos, de los cuales 4 son defectuosos. Si se eligen 5 base de datos al azar y sin reemplazo hallar la probabilidad de elegir no más de 2 base de datos defectuosos.

X=Número de base de datos defectuosos elegidos

N=10,   A=4,

n=5,   B=10-4=6

$$f(x) = \frac{\binom{4}{x}\binom{6}{5-x}}{\binom{10}{5}} \quad , \quad si \quad x = 0,1,2,3,4,5$$

$$= 0 \quad , \quad de\ otro\ modo$$

$$\mu_X = E(X) = \frac{nA}{N} = \frac{(5)(4)}{10} = 2 \quad \text{base de datos defectuosos}$$

$$\sigma_X^2 = Var(X) = \frac{nAB}{N^2}\left[\frac{N-n}{N-1}\right] = \frac{(5)(4)(6)}{10^2}\left[\frac{10-5}{10-1}\right] = 0.666667$$

Distribuciones discretas.

Distribución Hipergeométrica Ejemplo

Analytics AoZ

$$f(x) = \frac{\binom{4}{x}\binom{6}{5-x}}{\binom{10}{5}} \quad , \quad si \quad x = 0,1,2,3,4,5$$

$$= 0 \quad , \quad de\ otro\ modo$$

$$\mathrm{P}[X \leq 1] = f(0) + f(1) = \frac{\binom{4}{0}\binom{6}{5}}{\binom{10}{5}} + \frac{\binom{4}{1}\binom{6}{4}}{\binom{10}{5}} = 0.261905$$

DISTRIBUCIONES DE PROBABILIDAD

Analytics AoZ

Distribuciones continuas

Son las funciones de probabilidad asociadas a variables aleatorias continuas; es decir, para aquellas variables cuyo rango es un conjunto infinito no numerable.

En estos casos el valor de la función de probabilidad no expresa la probabilidad de ocurrencia de un valor específico de la variable aleatoria.

Ejemplos de vida real de distribuciones: https://www.somesolvedproblems.com/2019/04/real-life-examples-of-various.html#:~:text=Example%20of%20Uniform,roughly%201%2C000%20of%20each%20result.

Distribuciones continuas
Distribución Uniforme

Una variable aleatoria continua X tiene una distribución Uniforme si su función de probabilidad es dada por:

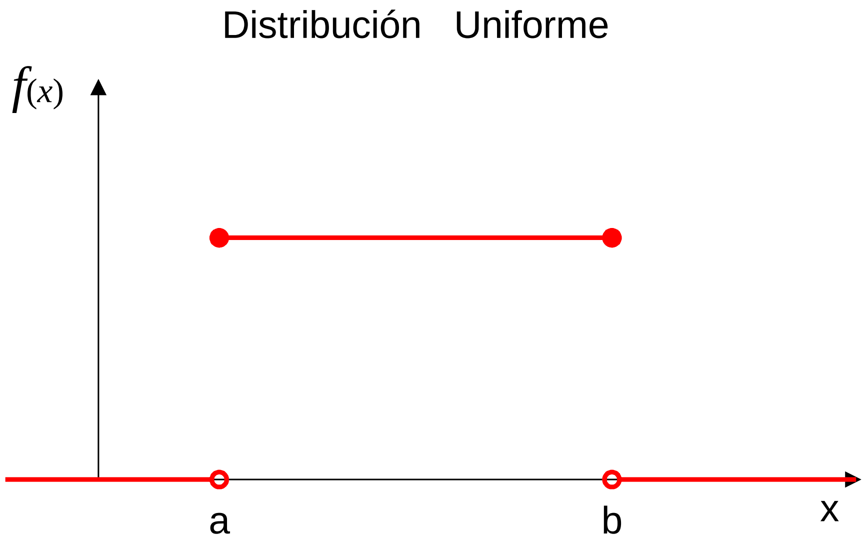
$$f(x) = \frac{1}{b-a} \quad , \quad si \quad a \le x \le b$$

$$= 0 \quad , \quad de\ otro\ modo$$

donde:

$$\mu_X = \mathrm{E}[X] = (a+b)/2$$

$$\sigma_X^2 = \mathrm{Var}[X] = {(b-a)^2}/{12}$$



Distribuciones continuas
Distribución Uniforme

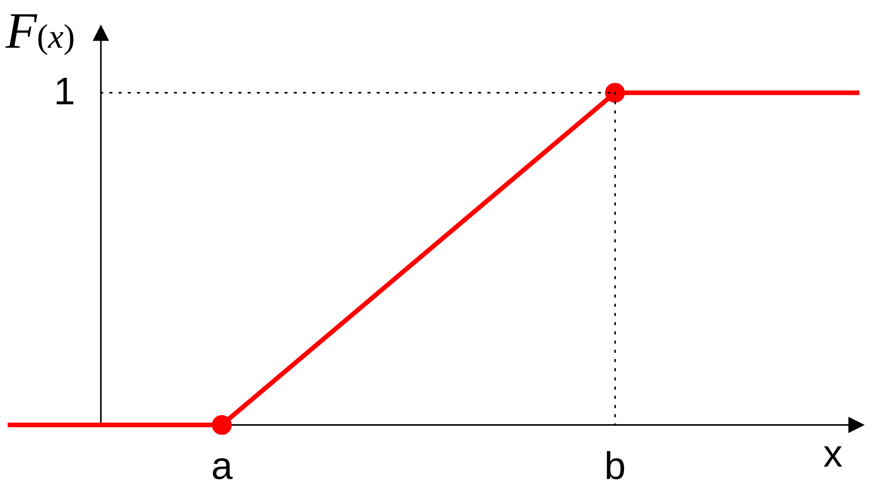
La función de distribución acumulativa de probabilidades es:

$$F(x) = 0 \quad , \quad si \quad x < a$$

$$= \frac{x-a}{b-a} \quad , \quad si \quad a \leq x \leq b$$

$$= 1 \quad , \quad si \quad x > b$$

Distribución acumulativa de probabilidades



Distribuciones continuas
Distribución Gamma

Una variable aleatoria continua X tiene una distribución Gamma si su función de probabilidad es dada por:

$$f(x) = \frac{1}{\beta^{\alpha}\,\Gamma(\alpha)}\, x^{\alpha-1} e^{-x/\beta} \quad , \quad \text{si } \; x > 0$$

$$= 0 \qquad\qquad\qquad , \quad \text{si } \; x \leq 0$$

donde: $\alpha>0$ **,** $\beta>0$ **y**

$$\mu_X = \mathrm{E}\lfloor X \rfloor = \alpha\beta$$

$$\sigma_X^2 = \mathrm{Var}[X] = \alpha\beta^2$$

$$\Gamma(\alpha) = \int_0^{\infty} e^{-t} t^{\alpha-1} dt, \quad si \;\; \alpha > 0$$

$$\Gamma(\alpha+1) = \alpha! \quad ,$$

$$\text{si} \quad \alpha = 0,1,2,3,\cdots$$

Distribuciones continuas
Distribución Gamma

Distribución Gamma

$$X \approx G(\alpha, \beta)$$

$$\mu_X = \alpha\beta$$

$$\sigma_X^2 = \alpha\beta^2$$

$$\alpha = 1$$

Distribución Exponencial

$$X \approx E(\beta)$$

$$\mu_X = \alpha\beta = \beta$$

$$\sigma_X^2 = \alpha\beta^2 = \beta^2$$

$$\alpha = n/2$$

$$\beta = 2$$

Distribución Chi-cuadrado

$$X \approx \chi^2(n)$$

$$\mu_X = \alpha\beta = n$$

$$\sigma_X^2 = \alpha\beta^2 = 2n$$

Distribuciones continuas
Distribución Exponencial

Una variable aleatoria continua X tiene una distribución Exponencial si su función de probabilidad es dada por:

$$
\begin{aligned}
f(x) &= \frac{1}{\beta} e^{-x/\beta} \quad , \text{ si } \quad x > 0 \\
&= 0 \qquad\qquad , \text{ si } \quad x \leq 0
\end{aligned}
$$

donde:

$$\mu_X = \mathrm{E}[X] = \beta$$

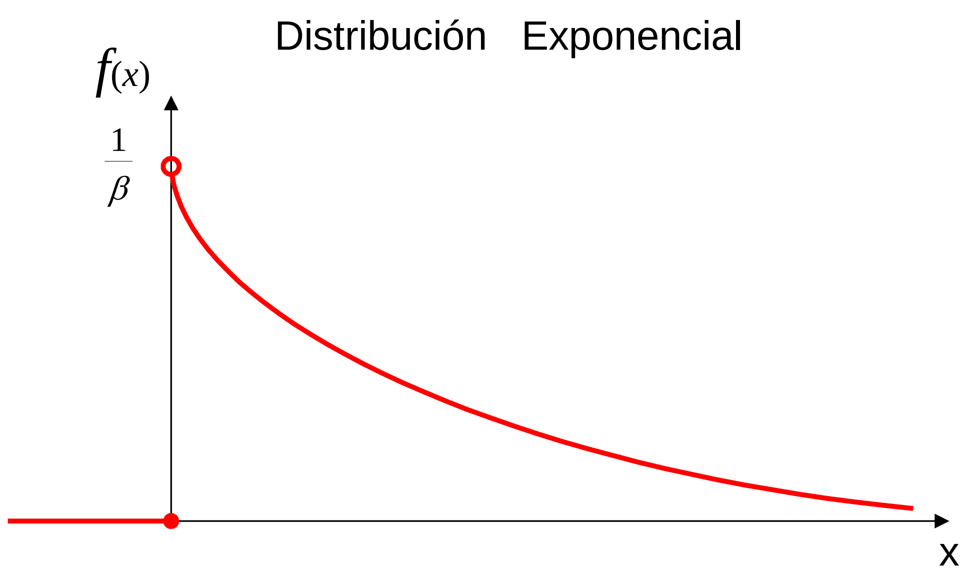
$$\sigma_X^2 = \mathrm{Var}[X] = \beta^2$$

Distribuciones continuas
Distribución Exponencial

Características

1. Esta definida para valores positivos de la variable.
2. Es asintótica al eje horizontal en su parte positiva.
3. Es monótona decreciente.

$$f(x) = \frac{1}{\beta} e^{-x/\beta} \quad , \text{ si } \quad x > 0$$
$$= 0 \quad , \text{ si } \quad x \leq 0$$

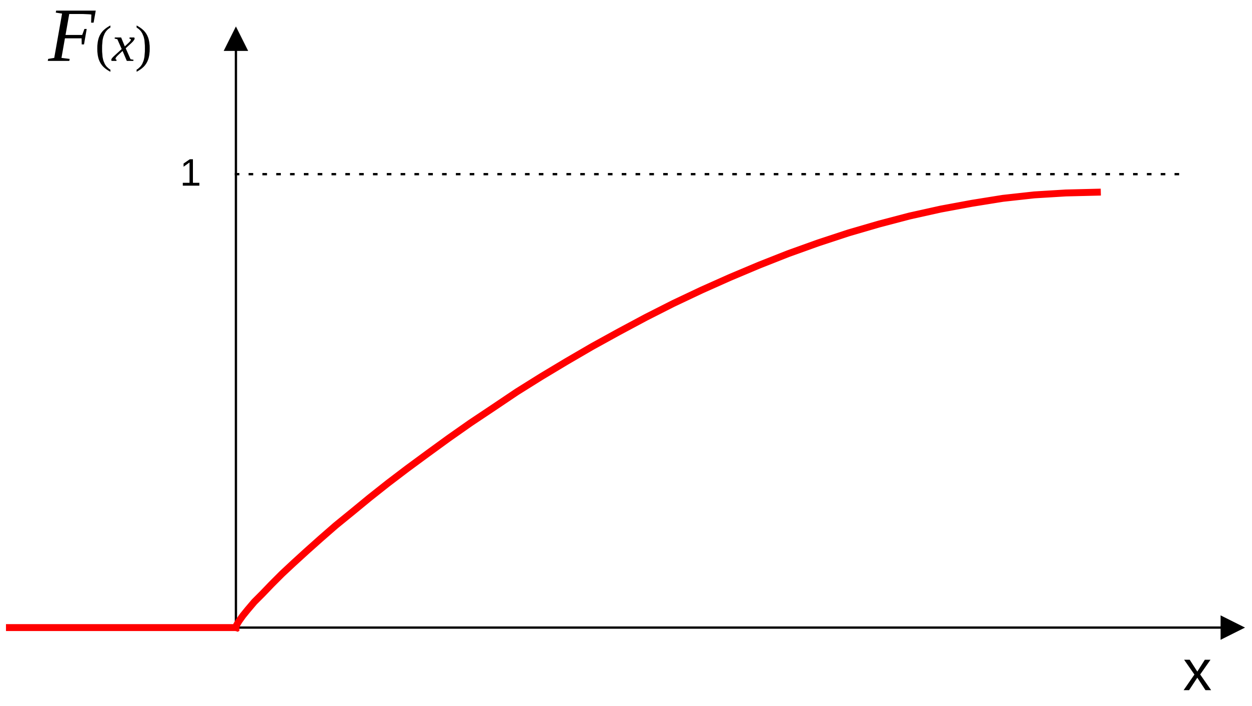


Distribuciones continuas
Distribución Exponencial

La función de distribución acumulativa de probabilidades es:

$$F(x) = 0 \qquad , \quad si \quad x \leq 0$$

$$= 1 - e^{-x/\beta}, \quad si \quad x > 0$$



Distribuciones continuas

Distribución Exponencial Ejemplo

Analytics AoZ

Suponga que el tiempo de duración de un artículo (horas) es una variable que tiene como densidad:

$$f(x) = \frac{1}{2} e^{-x/2} \quad , \text{ si } \quad x > 0$$

$$= 0 \qquad\qquad , \text{ si } \quad x \leq 0$$

Halle la probabilidad que un artículo elegido al azar tenga un tiempo de vida mayor a 3 horas.

$$\mathrm{P}[\mathrm{X} > 3] = \int_3^{\infty} f(x)dx = \lim_{a\to\infty} \int_3^{a} f(x)dx = \lim_{a\to\infty} \int_3^{a} \frac{1}{2} e^{-x/2} dx$$

$$= \lim_{a\to\infty} \left(- e^{-x/2} \right)\Big|_3^{a} = \lim_{a\to\infty} \left(- e^{-a/2} + e^{-3/2} \right) = e^{-3/2}$$

Distribuciones continuas
Distribución Normal

Una variable aleatoria continua X tiene una distribución Normal si su función de probabilidad es dada por:

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad \text{si } -\infty < x < \infty$$

donde:

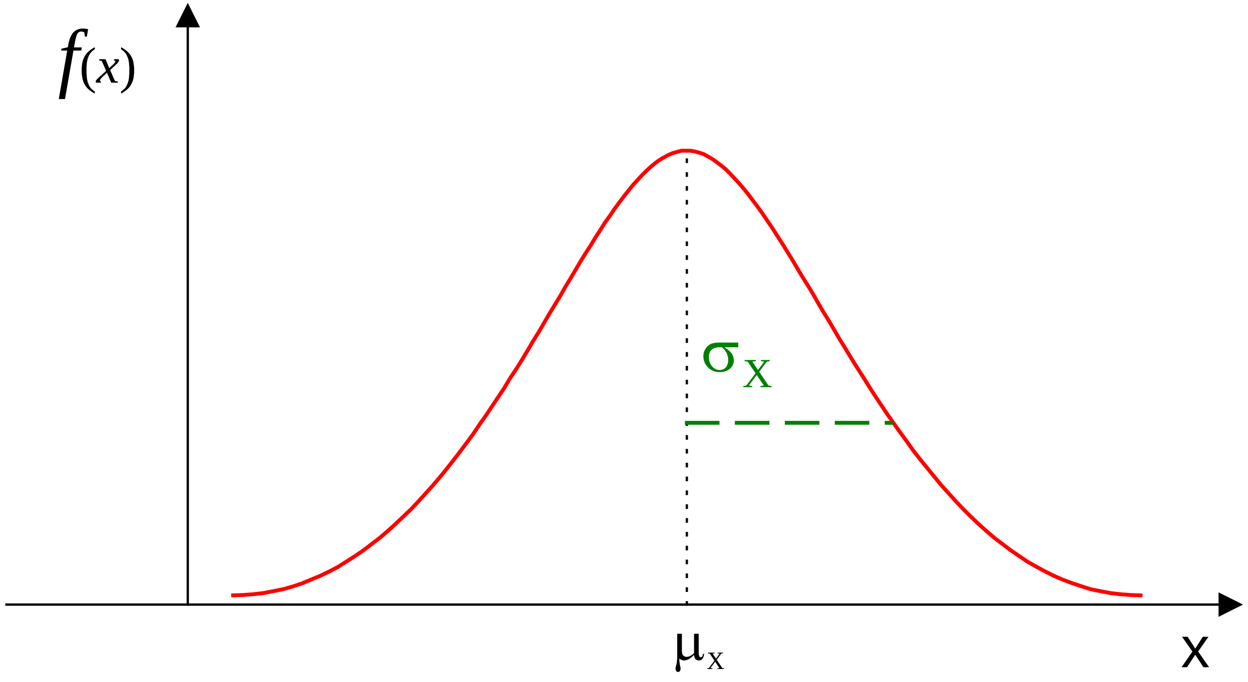
$$\mu_X = \mathrm{E}\lfloor X \rfloor = \mu$$

$$\sigma_X^2 = \mathrm{Var}[X] = \sigma^2$$

Distribuciones continuas
Distribución Normal

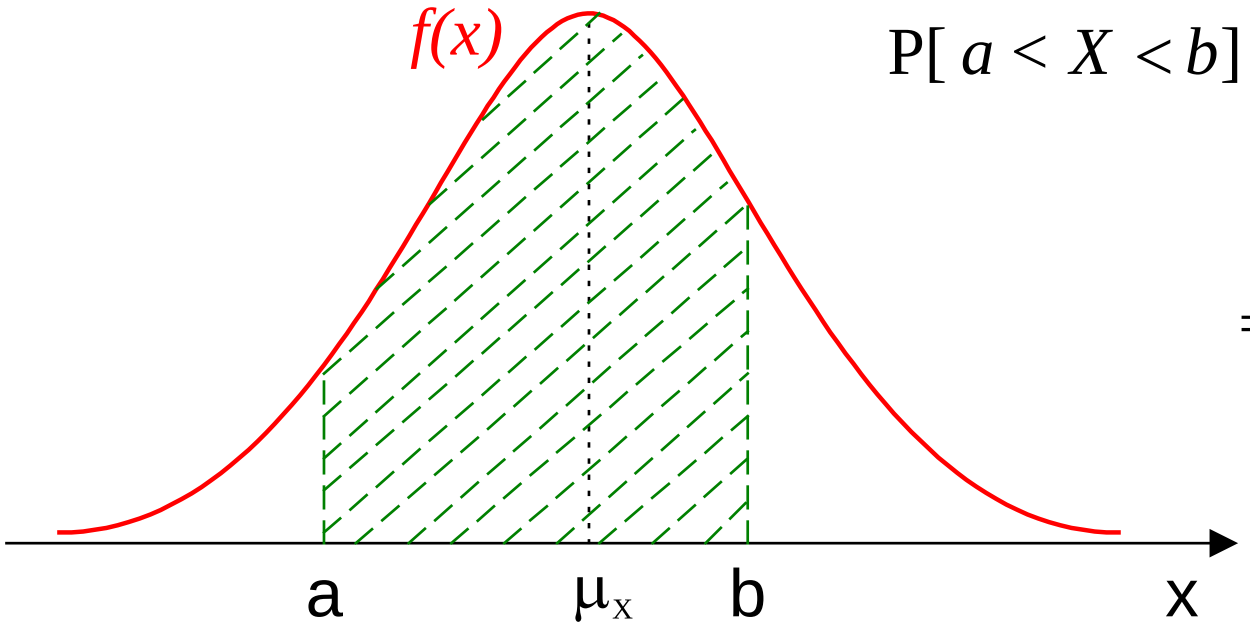
Características

1. Tiene una forma acampanada.
2. Es simétrica con respecto a si media.
3. Es asintótica con respecto al eje horizontal.



Distribuciones continuas
Distribución Normal

Cálculo de una probabilidad



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

$$= \int_a^b \frac{1}{\sqrt{2\pi}\,\sigma_X}\, e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2} dx$$

Si la distribución de los datos es simétrica con media y desviación estándar se cumplen las siguientes características empíricas:

- ✓ Aproximadamente, el 68% de los datos están comprendidos en el intervalo:

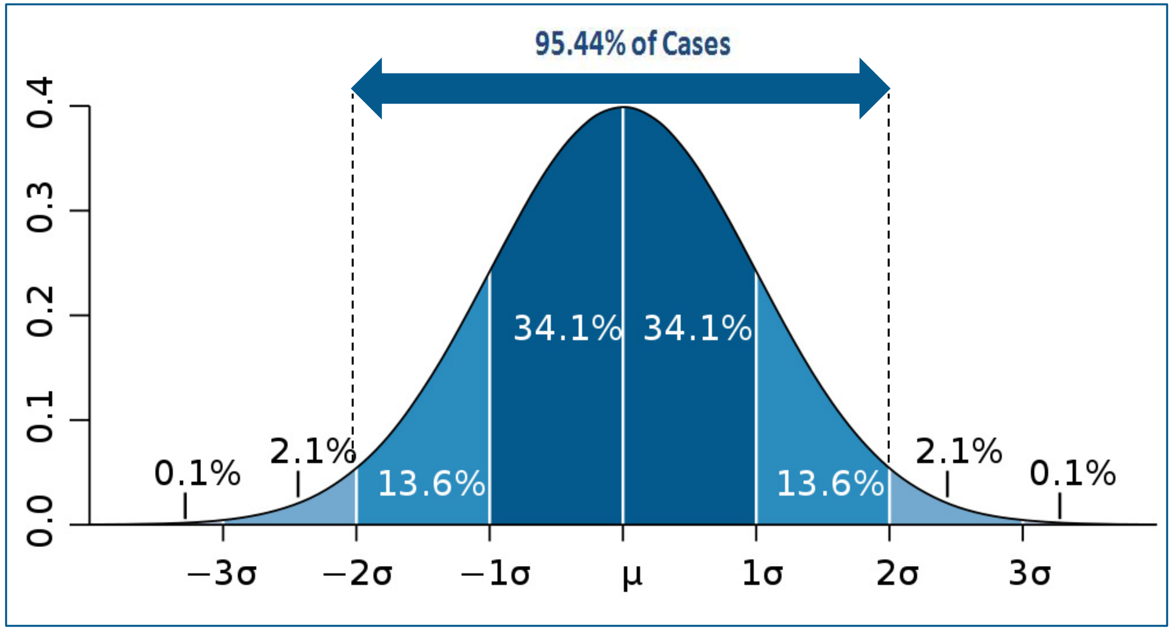
$$[\ \mu - \sigma,\ \ \mu + \sigma\ ]$$

- ✓ Aproximadamente, el 95% de los datos están comprendidos en el intervalo:

$$[\ \mu - 2\sigma,\ \ \mu + 2\sigma\ ]$$

- ✓ Aproximadamente, el 99% de los datos están comprendidos en el intervalo:

$$[\ \mu - 3\sigma,\ \ \mu + 3\sigma\ ]$$



Independiente de la forma de simetría de los datos: se puede afirmar que el intervalo $[\ \mu - k\sigma,\ \ \mu + k\sigma\ ]$, donde k > 1, contiene por lo menos $\left(1 - \frac{1}{k^2}\right)\%$ datos.

DISTRIBUCIÓN NORMAL

Es una distribución de probabilidad especial para variables aleatorias continuas y posee ciertas propiedades importantes que conviene destacar:

- Tiene una única moda, que coincide con su media y su mediana.
- Es simétrica con respecto a su media; según esto, para este tipo de variables existe una probabilidad de un 50% de observar un dato mayor y menor que la media.
- **Muchas mediciones de diversos procesos se ajustan a esta distribución**.
- Se utiliza para ajustar variables que tienen otros tipo de distribución como *Binomial, Poisson, T-Student, Chi-Cuadrado, F-Fisher*, etc.
- Las estadísticas muestrales como la media y la proporción tienen distribuciones que se ajustan a es tipo de distribución cuando el tamaño de muestra es grande, inclusive de aquellas que no tiene distribución normal.

DISTRIBUCIÓN NORMAL

Analytics AoZ

FUNCIÓN DE DISTRIBUCIÓN DE PROBABILIDAD

La función de distribución de probabilidad de la Normal, se puede interpretar también como la distribución acumulada de probabilidad de una variable X que se distribuye como una normal y está dada por:

$$F(x) = P[X \leq x] = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

La ***Media*** y la ***Varianza*** de una variable con Distribución Normal están definidas de la siguiente forma:

$$E(X) = \mu \quad y \quad V(X) = \sigma^2$$

Propiedades

Si $X \sim N(\mu X, \sigma X)$ e $Y \sim N(\mu Y, \sigma Y)$ son independientes, entonces la distribución de la suma o diferencia de ambas variables es también normal con las siguientes medias y desviaciones típicas.

$$X + Y \sim N\left(\mu_X + \mu_Y, \sqrt{\sigma_X^2 + \sigma_Y^2}\right)$$

$$X - Y \sim N\left(\mu_X - \mu_Y, \sqrt{\sigma_X^2 + \sigma_Y^2}\right)$$

Ejemplo 1

El tiempo necesario para terminar el examen parcial de un determinado curso se distribuye normalmente con un tiempo promedio de 80 minutos y una desviación absoluta de 10 minutos.

a) ¿Cuál es la probabilidad de que un alumno termine el examen en más de 60 minutos pero en menos de 70 minutos?

b) Suponga que en el grupo hay 60 alumnos y que el tiempo del examen es de 90 minutos. ¿Cuántos alumnos se debe esperar que no puedan terminar el examen en el tiempo indicado?

EJEMPLO

En determinado establecimiento se vende tres marcas diferentes de coches. Sean $X1$, $X2$ y $X3$ variables aleatorias independientes normales, que representan el volumen semanal de ventas para cada una de las marcas.

Las ventas medias semanales de estas marcas son 42, 60 y 78 mil euros, respectivamente, y sus desviaciones típicas respectivas son 12, 18 y 10 mil euros.

EJEMPLO

a) *Cuál es la probabilidad de que la primera marca no supere los 30 mil euros en una semana?. Y la probabilidad de que la segunda marca supere en un semana la mediana de la tercera marca?*

a) *Calcular la probabilidad de que, en una semana determinada, las ventas del establecimiento sean superiores a los 120 mil euros.*

b) *Cuál es la probabilidad de que la suma de las ventas de la primera marca y de la tercera superen a las ventas de la segunda marca en mas de 18 mil euros, en una semana?*

Analytics AoZ

Distribuciones continuas
Distribución Normal Estándar

Si X es una variable aleatoria que tiene una distribución normal con una media μ_X y una variancia σ_X^2 ; luego, la variable aleatoria Z = (X - μ_X)/ σ_X tiene una distribución normal con media 0 y variancia 1, y su función de densidad es:

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

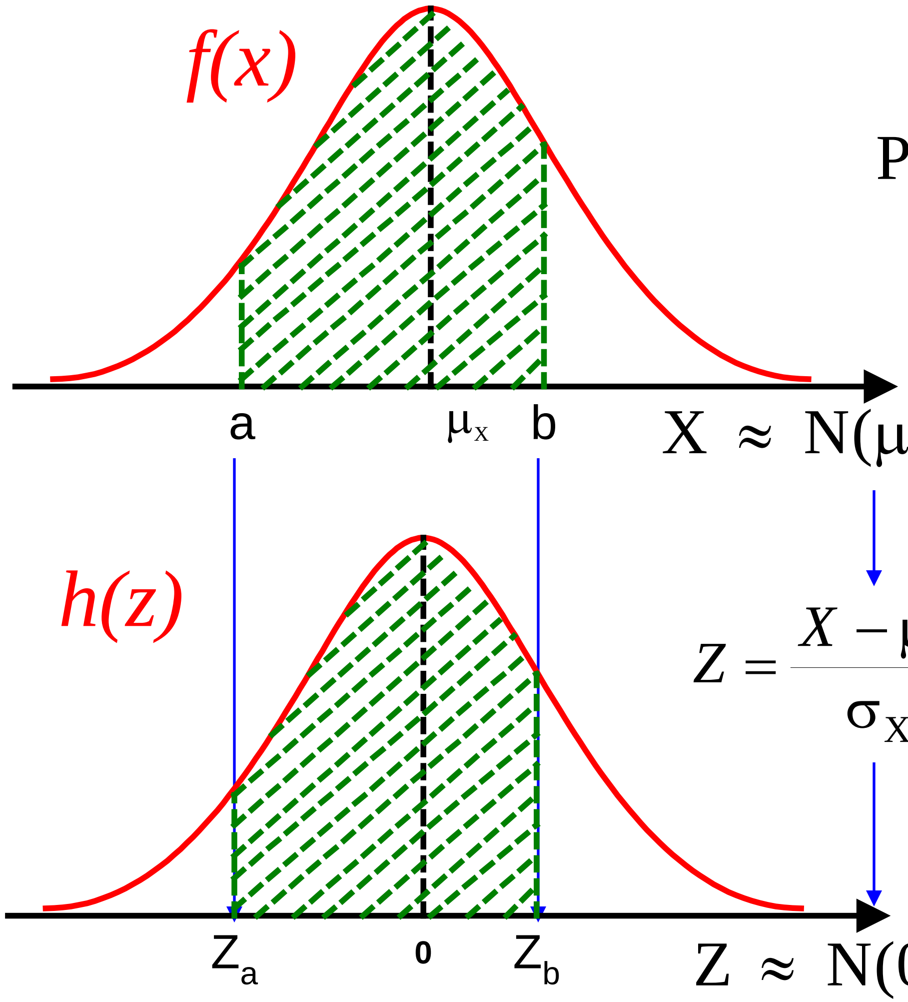
donde:

$$\mu_Z = \mathrm{E}[Z] = 0$$

$$\sigma_Z^2 = \mathrm{Var}[Z] = 1$$

Distribuciones continuas

Distribución normal estándar. Cálculo de una probabilidad



$$X \approx N(\mu_X, \sigma_X^2)$$

$$Z = \frac{X - \mu_X}{\sigma_X}$$

$$Z \approx N(0,1)$$

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

$$= \int_a^b \frac{1}{\sqrt{2\pi}\,\sigma_X} e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2} dx$$

$$= \int_{Z_a}^{Z_b} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

$$= P[\, z_a < Z < z_b]$$

Distribuciones continuas

Distribución normal estándar. Ejemplo

Suponga que los gastos semanales en movilidad realizados por los habitantes de un distrito tienen una distribución normal con una media de 50 nuevos soles y una desviación de 12 nuevos soles. Si se elige al azar un habitante de dicho distrito, halle la probabilidad que su gasto semanal en movilidad sea mayor de 58 nuevos soles.

X= Gasto semanal en movilidad (nuevos soles)

$$X \approx N\left[50, (12)^2\right]; \qquad \mu_X = 50, \qquad \sigma_X^2 = (12)^2$$

$$P\left[X > 58\right] = ?$$

Distribuciones continuas
Distribución normal estándar.

X= Gasto semanal en movilidad (nuevos soles)

$$X \approx N\left[50, (12)^2\right]; \qquad \mu_X = 50, \qquad \sigma_X^2 = (12)^2$$

$$P\left[X > 58\right] = ?$$

$$Z = \frac{X - \mu_X}{\sigma_X} \approx N\left[0, 1\right]; \qquad \mu_Z = 0, \qquad \sigma_Z^2 = 1$$

$$P\left[X > 58\right] = P\left[Z > \frac{58 - \mu_X}{\sigma_X}\right] = P\left[Z > \frac{58 - 50}{12}\right] = P\left[Z > 0.67\right]$$

$$= 1 - P\left[Z \leq 0.67\right]$$

$$= 1 - 0.7486 = 0.2514$$

Distribucion LogNormal
Distribución lognormal

When are Lognormal Distributions Used?

The most commonly used (and the most familiar) distribution in science is the normal distribution. The familiar "bell curve" models many natural phenomenon, from the simple (weights or heights) to the more complex. For example, the following phenomenon can all be modeled with a lognormal distribution:

- Milk production by cows.
- Lives of industrial units with failure modes that are characterized by fatigue-stress.
- Amounts of rainfall.
- Size distributions of rainfall droplets.
- The volume of gas in a petroleum reserve.

Many more phenomenon can be modeled with the lognormal distribution, such as the length of latent periods of infectious disease or species abundance[1].

Revisar: https://www.statisticshowto.com/lognormal-distribution/#:~:text=The%20familiar%20%E2%80%9Cbell%20curve%E2%80%9D%20models,are%20characterized%20by%20fatigue%2Dstress.
https://academic.oup.com/bioscience/article/51/5/341/243981
https://www.sciencedirect.com/topics/mathematics/lognormal-distribution

DISTRIBUCIONES ESPECIALES

Analytics AoZ

Por lo general, cuando las muestras son grandes, la distribución de probabilidad más usada es la distribución normal. Sin embargo, cuando las muestras son pequeñas la distribución de probabilidad difiere de caso en caso, y muchas veces no son normales.

Existen tres distribuciones de probabilidad que a menudo son usadas:

- Distribución **Chi-Cuadrado**,
- Distribución **F de Fish**er (o Snedecor) y
- Distribución **T-Student**.

DISTRIBUCIÓN CHI-CUADRADO

Analytics AoZ

Si $X_1, \ldots, X_n$ son variables aleatorias independientes con distribución normal estándar, $X_i \sim$ ***N(0,1). Entonces la variable aleatoria***

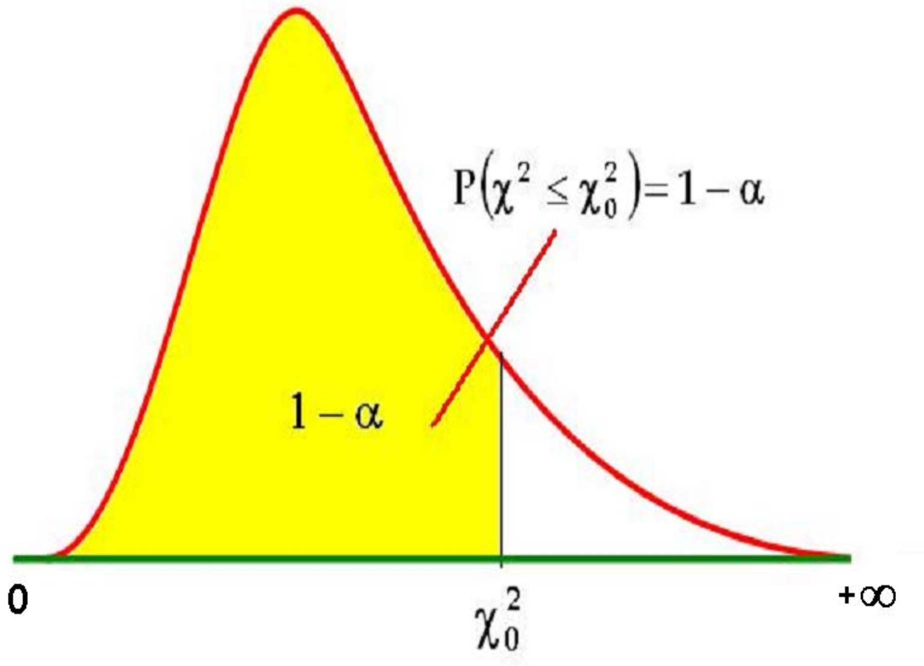
$$Y = X_1^2 + \ldots + X_n^2$$

Es decir, Y tiene una Distribución Chi-Cuadrado con n grados de libertad.

Esto es: $Y \sim \chi_n^2$

Si la función de densidad de Y está dada por:

$$f(Y) = \chi_n^2 = \frac{Y^{\frac{n}{2}-1} \quad e^{-\frac{Y^2}{2}}}{2^{\frac{n}{2}} \quad \Gamma(n/2)} \quad \forall \quad 0 \leq Y \leq \infty$$



El valor esperado y la varianza de esta variable están dados por:

E(Y) = n

V(Y) = 2n.

Ejemplo:

Si X es una variable que se distribuye como una distribución Chi-Cuadrado de 12 grados de libertad. Entonces se pide hallar los valores de a y b, tales que $P(a \leq X \leq b) = 0{,}90$ y $P(X \leq a) = 0{,}05$.

Solución .-

Si $X \sim \chi^2_{(12)}$ luego sabiendo que:

$$P(a \leq X \leq b) = 0{,}90$$

$$P(X \leq b) - P(X \leq a) = 0{,}90$$

$$\text{Como: } P(X \leq a) = 0{,}05$$

$$P(X \leq b) - 0{,}05 = 0{,}90$$

$$P(X \leq b) = 0{,}95$$

DISTRIBUCIÓN T-STUDENT

Si X_0, X_1, ..., X_n son variables aleatorias con distribución normal estándar, se puede afirmar que:

$$V = \frac{X_0}{\sqrt{\dfrac{X_1^2 + \ldots + X_n^2}{n}}} = \frac{X_0}{\sqrt{\dfrac{1}{n}\chi_n^2}}$$

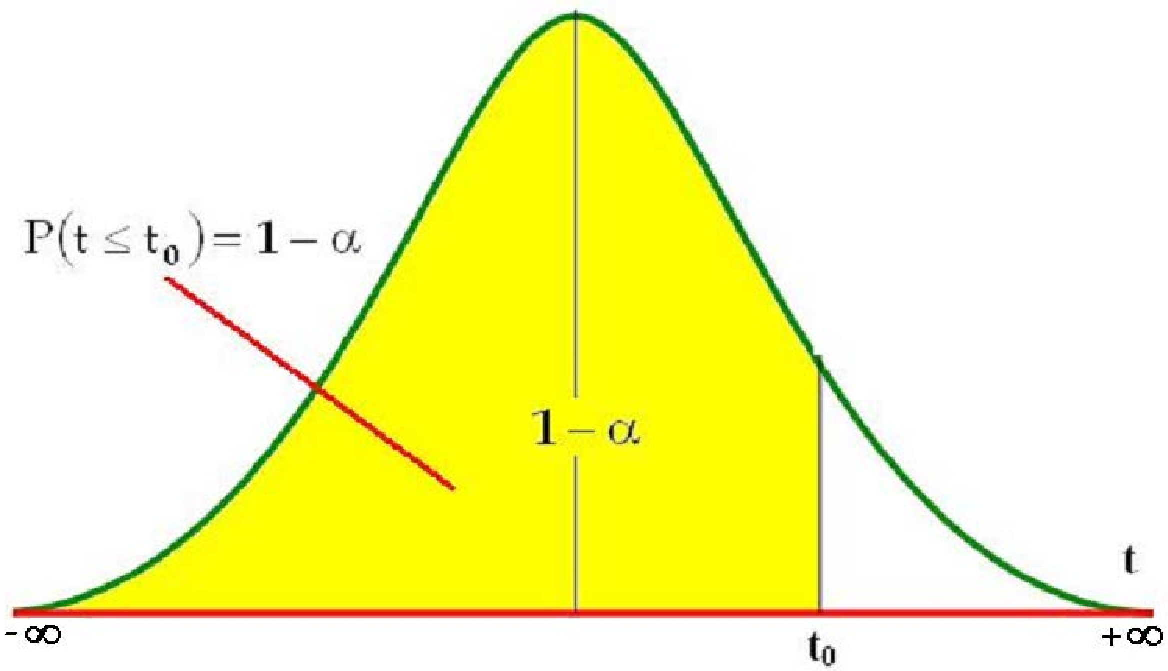
Es decir, Y tiene una Distribución T-Student con n grados de libertad, esto es, $t \sim t_{(n)}$.

La función de densidad de Y está dada por:

$$f(V) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\left(n+1\right)/2\right)}{\Gamma\left(n/2\right)} \left(1 + \frac{1}{n}Y^2\right)^{-\frac{n+1}{2}} \quad \forall \quad -\infty \leq Y \leq \infty$$

DISTRIBUCIÓN T-STUDENT

La grafica de la función de densidad tiene la siguiente forma:



El valor esperado y la varianza de esta variable están dados por:

$$E(V) = 0, \forall \ \ n > 1 \qquad V(V) = \frac{n}{n-2}, \ \ \forall \ \ n > 2$$

Ejemplo:

Analytics AoZ

Si X es una variable que tiene distribución t de Student con una varianza 5/4. Calcule: P(-1,812 ≤ X ≤ 2,228)

Solución

$$\text{Se sabe que:} \quad V(X) = \frac{n}{n-2} = \frac{5}{4}$$

$$\text{Luego:} \quad 4n = 5(n-2)$$

$$\text{Entonces:} \quad n = 10$$

Así $X \sim t_{(10)}$ y por tanto: $P(-1{,}812 \leq X \leq 2{,}228) = 0{,}9250$

DISTRIBUCIÓN F DE FISHER

Si $X_1, \ldots, X_n$ son variables aleatorias con distribución normal estándar; y además se tiene que $Y_1, \ldots, Y_m$ son variables aleatorias que también tienen distribución normal estándar, se puede afirmar que:

$$W = \frac{(X_1^2 + \ldots + X_n^2)/n}{(Y_1^2 + \ldots + Y_m^2)/m} = \frac{\chi_n^2/n}{\chi_m^2/m}$$

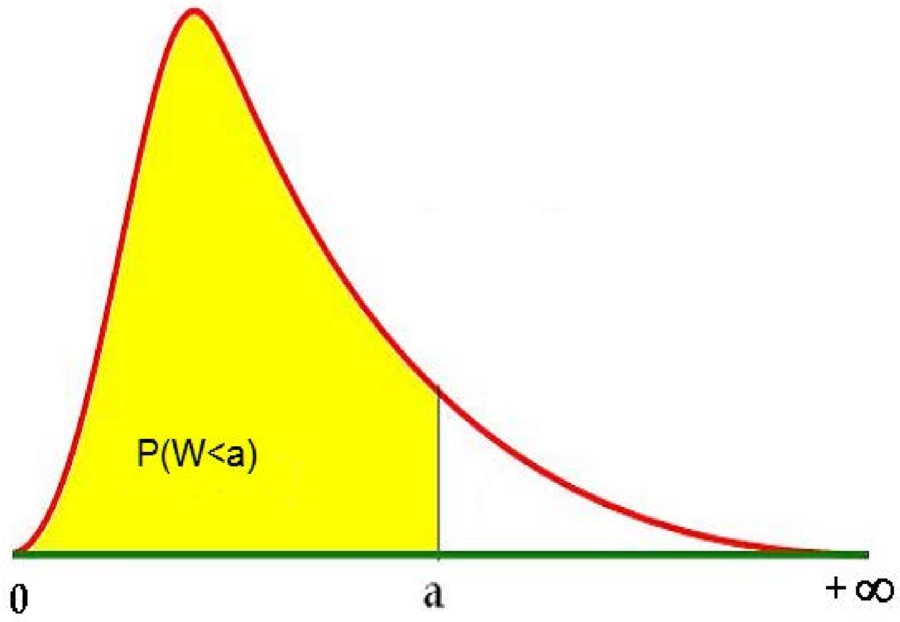
Es decir, W tiene una Distribución F con n y m grados de libertad. Esto es: $W \sim F_{(n;\, m)}$

La función de densidad de W está dada por:

$$f(W) = \frac{\Gamma\left((n+m)/2\right)}{\Gamma(n/2)\ \Gamma(m/2)} \left(\frac{m}{n}\right)^{\frac{m}{2}} W^{\frac{m}{2}-1} \left(1+\frac{m}{n}W\right)^{-\frac{m+n}{2}}, \ \forall \ \ 0 \le W \le \infty$$

DISTRIBUCIÓN F DE FISHER

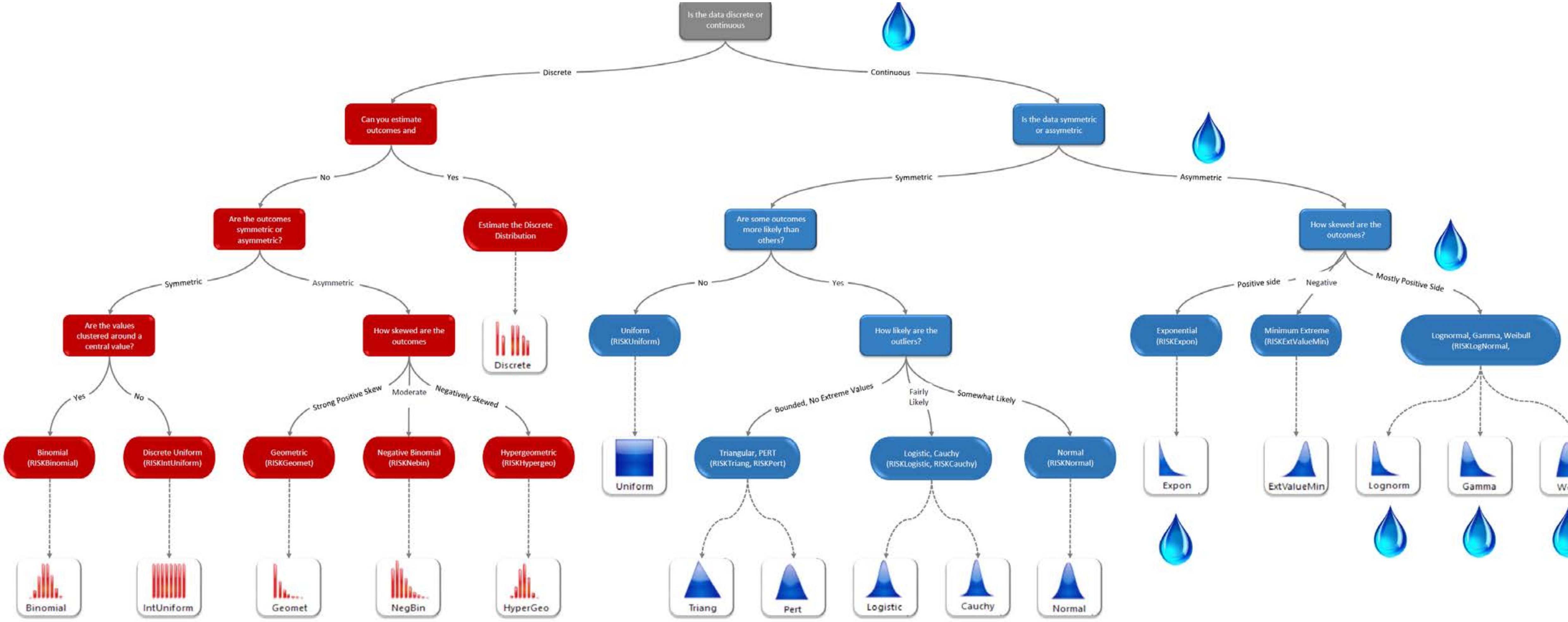
La grafica de la función de densidad tiene la siguiente forma:



El valor esperado y la varianza de esta variable están dados por:

$$E(W) = \frac{n}{n-2}, \forall \quad n > 2 \qquad V(W) = \frac{2n^2(n+m-2)}{m(n-2)^2(n-4)}, \quad \forall \quad n > 4$$

¿Cómo se distribuye tu data?



Fuente: https://www.palisade.com/

Consejos

- Las variables aleatorias deben estar bien definidas en nuestro conjunto de datos tanto sean continuas o discretas ***debemos reconocerlas.***
- En los problemas de “modelado” o análisis debemos testear que tipo de distribución puede poseer nuestra data (las variables) ***pensemos en variedad.***
- Es una recomendación muy importante resumir y visualizar los datos de manera gráfica y numéricamente para testear el tipo de distribución de nuestros datos, **probar diferentes técnicas**.

EL AGUA

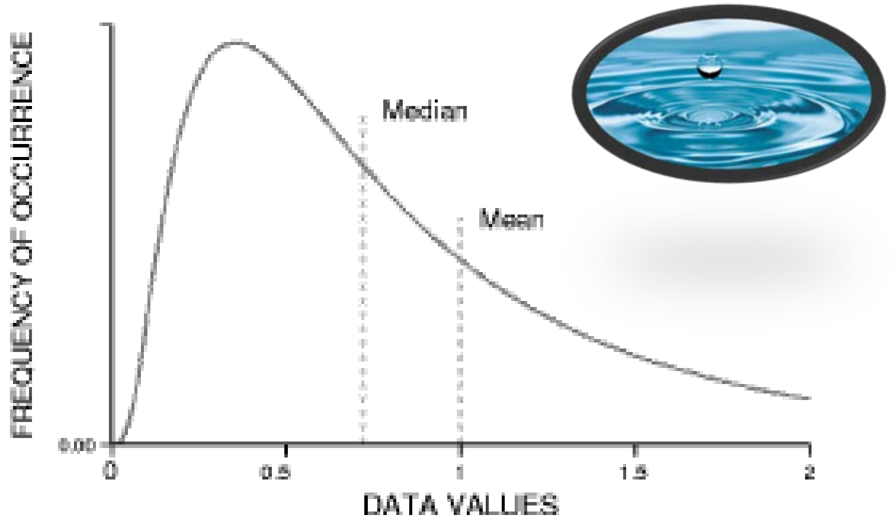


Figure 1.1 Density Function for a Lognormal Distribution

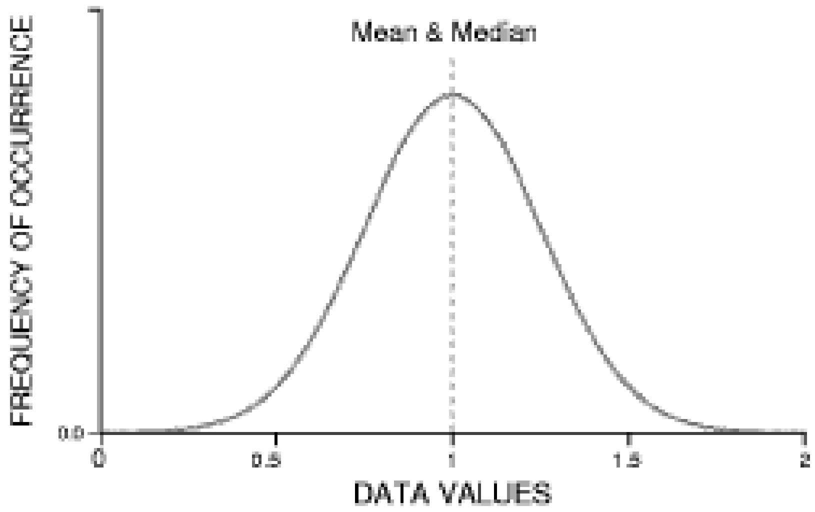
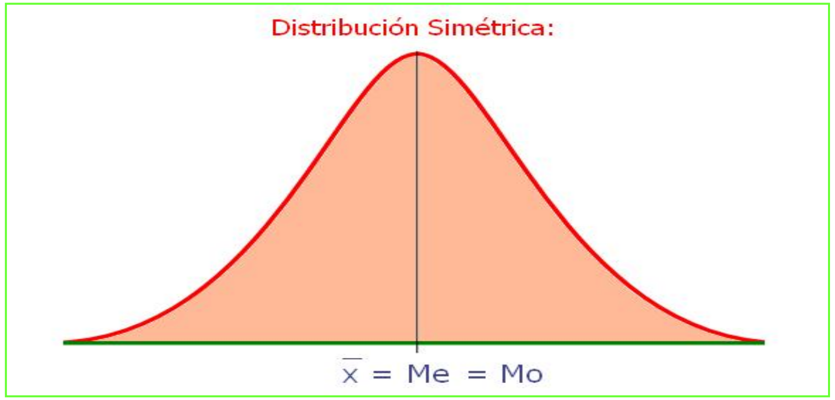


Figure 1.2 Density Function for a Normal Distribution

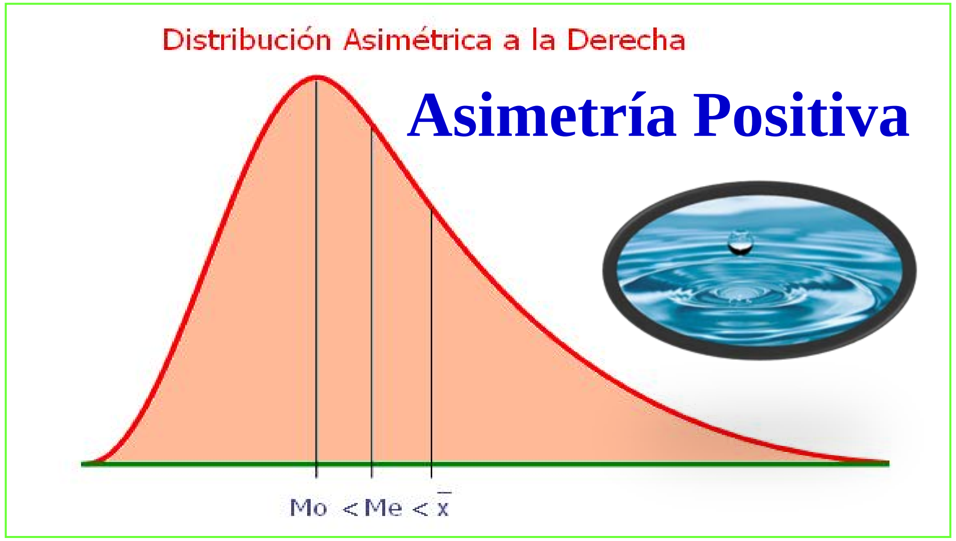
Relación Media, Mediana y Moda

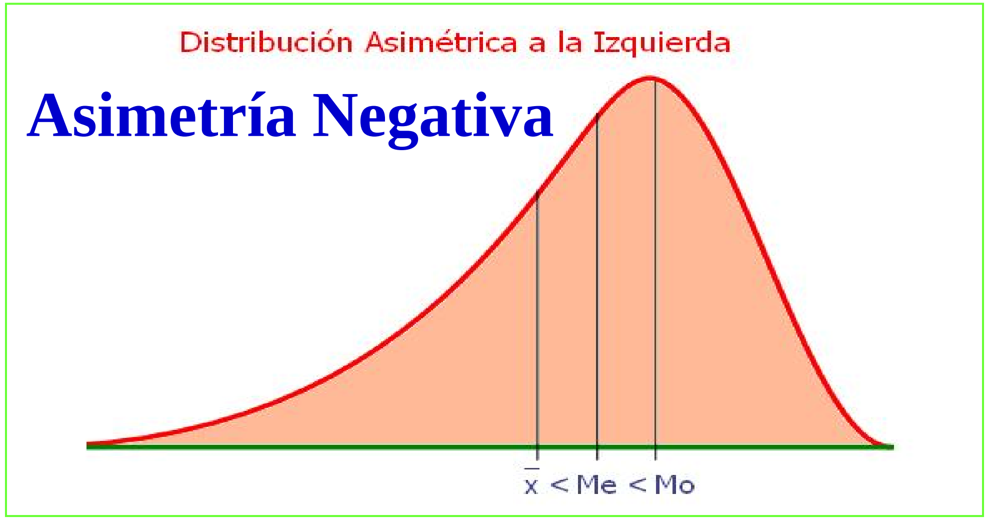
En distribuciones uni-modales se cumple:

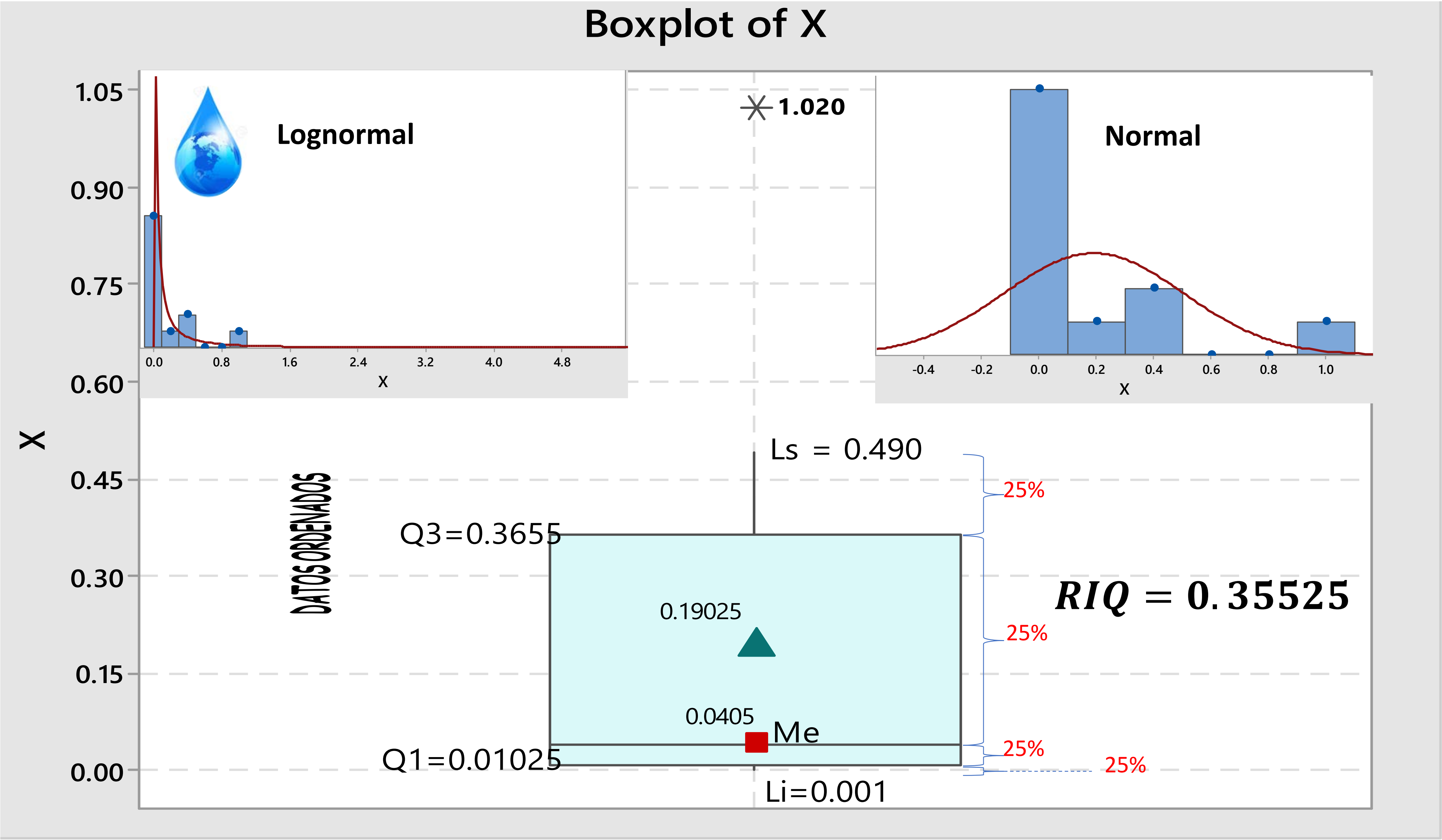
- Si una distribución es simétrica, la media, mediana y moda coinciden



- Si una distribución no es simétrica, las tres medidas difieren.

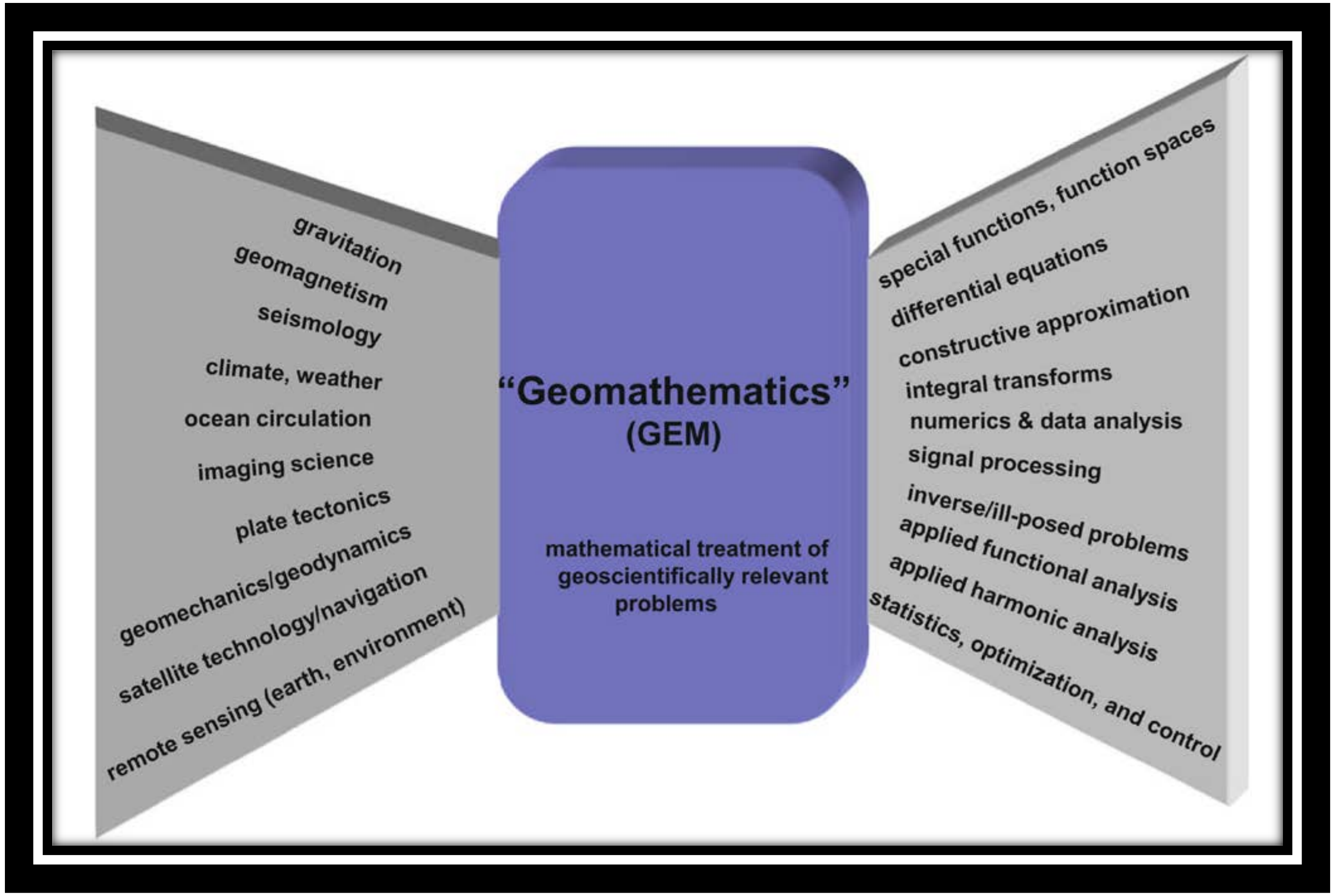


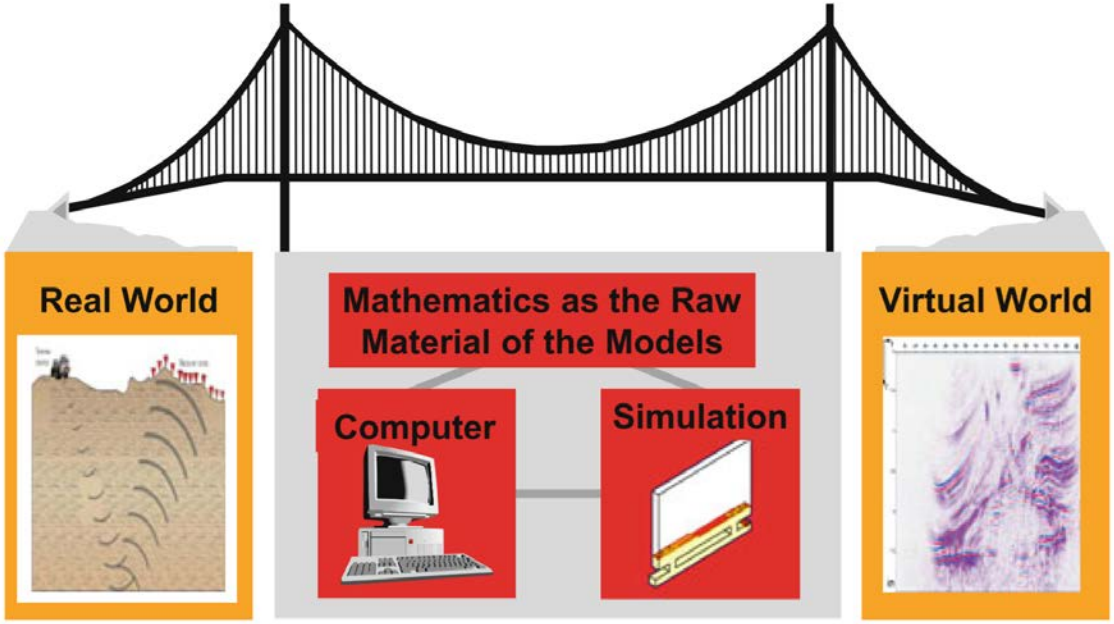




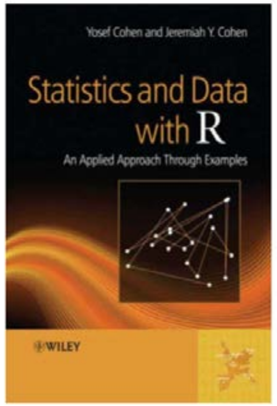
Boxplot of X
Lognormal
Normal
1.020
Ls = 0.490
Q3=0.3655
0.19025
0.0405
Me
Q1=0.01025
Li=0.001
DATOS ORDENADOS
RIQ = 0.35525
25%
25%
25%
25%
X
1.05
0.90
0.75
0.60
0.45
0.30
0.15
0.00

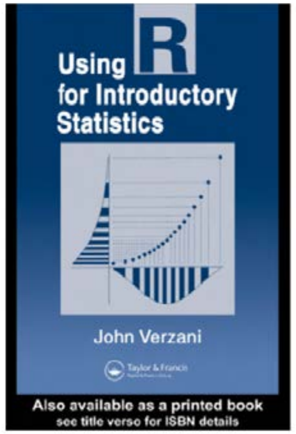
CENTER OF GEOMATHEMATICS OF ENGINEERING SCHOOL

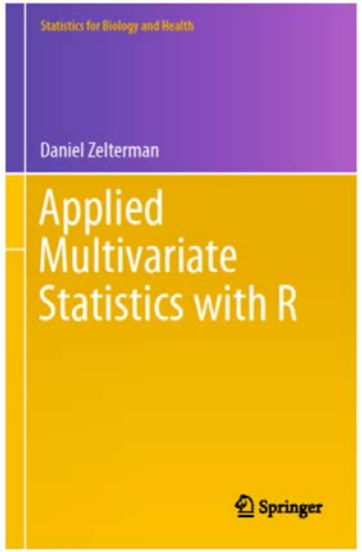


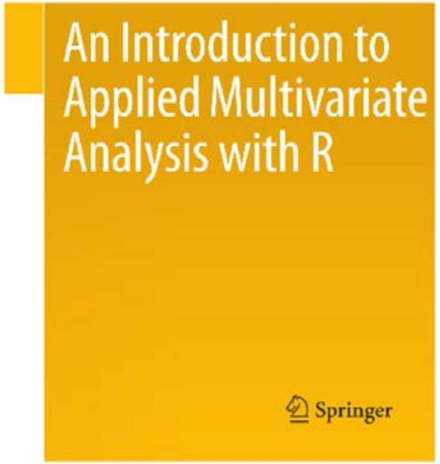


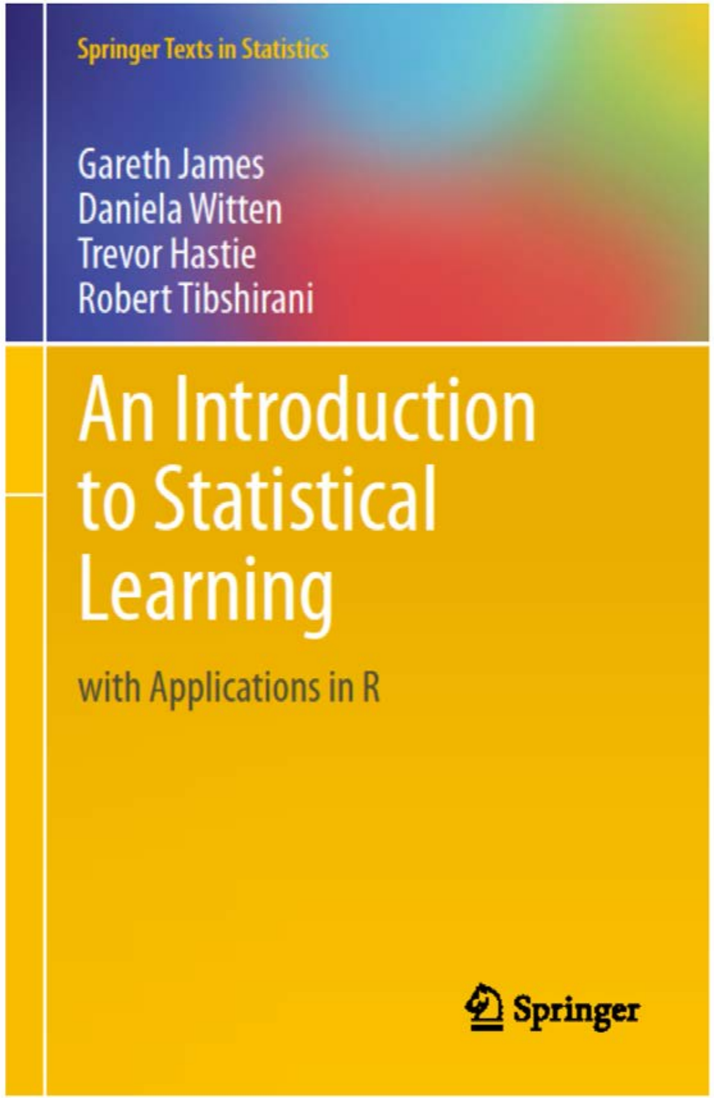
BIBLIOGRAFÍA:

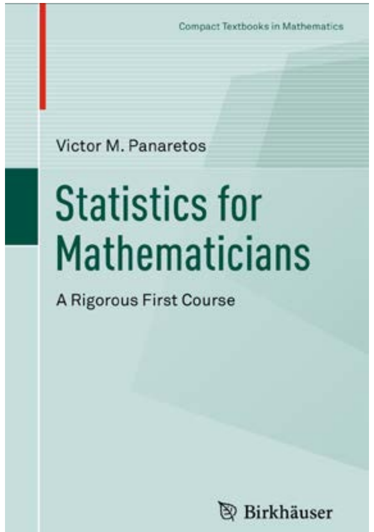


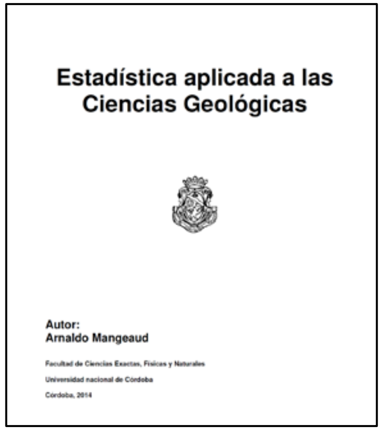


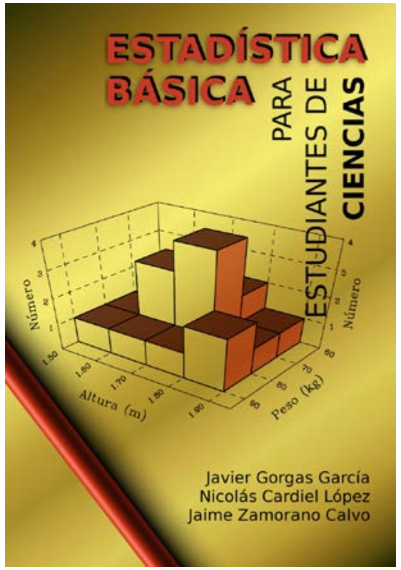












BIBLIOGRAFÍA ESTADÍSTICA APLICADA A GEOLOGÍA EXTRA:

Practical Methods for Data Analysis (US EPA QA/G-9, 2000)

Helsel, D. R., & Hirsch, R. M. (2002). *Statistical methods in water resources* (Vol. 323). Reston, VA: US Geological Survey.

Salvador Figu
eras, M y Gargallo, P. (2003): "Análisis Exploratorio de Datos, 5campus.com, Estadística <http://www.5campus.com/leccion/aed>

Ramalle-Gómara, E., & De Llano, J. A. (2003). Utilización de métodos robustos en la estadística inferencial. *Atención Primaria*, *32*(3), 177-182.

Verzani, J. (2005). *Using R for introductory statistics*. CRC press.

Cohen, Y., & Cohen, J. Y. (2008). *Statistics and Data with R: An applied approach through examples*. John Wiley & Sons.

Arnaldo Mangeaud (2014). *Estadística aplicada a las Ciencias Geológicas*. Universidad nacional de Córdova.

Helsel, D.R., Hirsch, R.M., Ryberg, K.R., Archfield, S.A., and Gilroy, E.J., 2020, *Statistical methods in water resources: U.S. Geological Survey Techniques and Methods*, book 4, chapter A3, 458 p.

Analytics AoZ

“LO QUE ESCUCHO LO OLVIDO. LO QUE VEO LO RECUERDO. PERO LO QUE HAGO, LO ENTIENDO.”

AUTOR: ANÓNIMO

VAMOS A RESOLVER UN EJERCICIO