

# **EYP1113 - PROBABILIDAD Y ESTADÍSTICA**

## **CAPÍTULO 3: MODELOS ANALÍTICOS DE FENÓMENOS ALEATORIOS**

**RICARDO ARAVENA C.**

**RICARDO OLEA O.**

**FACULTAD DE MATEMÁTICAS**

**DEPARTAMENTO DE ESTADÍSTICA**

**PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE**

**SEGUNDO SEMESTRE 2019**

# CONTENIDO I

- 1 Variables Aleatorias y Distribución de Probabilidad
  - Eventos Aleatorios y Variables Aleatorias
  - Distribución de Probabilidades de una Variable Aleatoria
  - Medidas Descriptivas de Variables Aleatorias
- 2 Distribuciones de Probabilidad
  - Distribución Normal
  - Distribución Log-Normal
  - Distribución Binomial y Bernoulli
  - Distribución Geométrica
  - Distribución Binomial Negativa
  - Distribución de Poisson
  - Distribución Exponencial
  - Distribución Gamma
  - Distribución Hipergeométrica
  - Distribución Beta
- 3 Otras Distribuciones de Probabilidad

# CONTENIDO II

- Distribución Weibull
- Distribución Logística
- Distribución Log-Logística
- Distribución t-Student
- Distribución Fisher

## 4 Múltiples Variables Aleatorias

- Distribución de Probabilidad Conjunta y Condicional
- Distribuciones Marginales y Condicionales
- Covarianza y Correlación

## 5 Esperanza Condicional y Predicción

- Definiciones y Ejemplos
- Predicción

# **VARIABLES ALEATORIAS Y DISTRIBUCIÓN DE PROBABILIDAD**

Ya vimos los conceptos fundamentales y las herramientas referidas a la ocurrencia de fenómenos aleatorios y la determinación de las probabilidades asociadas.

Ahora, introduciremos modelos analíticos que involucran:

**Variables Aleatorias y Distribución de Probabilidad**

Una variable aleatoria es el vehículo matemático para representar un evento en términos analíticos.

El valor de una variable aleatoria puede estar definida para un conjunto de posibles valores.

Si  $X$  es una variable aleatoria, entonces

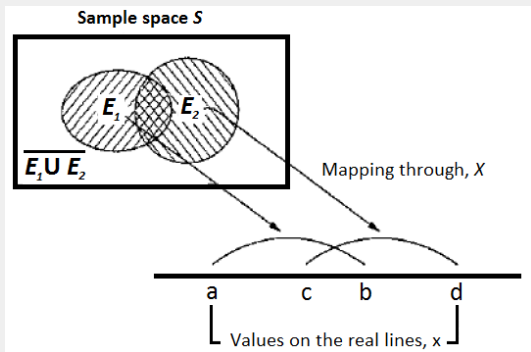
$$X = x, \quad X < x, \quad X > x$$

representa un evento, donde  $(a < X < b)$  es el rango de valores posibles de  $X$ .

La asignación numérica puede ser natural o artificial.

# VARIABLES ALEATORIAS Y DISTRIBUCIÓN DE PROBABILIDAD

Formalmente, una variable aleatoria puede ser considerada como una función o regla sobre los eventos del espacio muestral a un sistema numérico (o línea real).



Así, los eventos  $E_1$  y  $E_2$  pueden corresponder a

$$E_1 = (a < X \leq b)$$

$$E_2 = (c < X \leq d)$$

$$\overline{E_1 \cup E_2} = (X \leq a) \cup (X > d)$$

$$E_1 \cap E_2 = (c < X \leq b)$$

Una variable aleatoria puede ser discreta o continua.



Para los valores o rango de valores que puede tomar una variable aleatoria tienen asociados una probabilidad específica o medidas de probabilidad.

La regla que asigna las medidas de probabilidad se denomina:

**“Distribución ó Ley de probabilidad”**

Si  $X$  es variable aleatoria, la distribución de probabilidad puede ser descrita por su función de distribución de probabilidad acumulada denotada por:

$$F_X(x) = P(X \leq x) \quad \text{para todo } x \in \mathbb{R}$$

Si  $X$  es variable aleatoria discreta, entonces esta función puede ser expresada a través de la función de probabilidad “puntual” denotada por

$$p_X(x) = P(X = x)$$

Así,

$$F_X(x) = \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i \leq x} p_X(x_i),$$

con  $x_i \in \Theta_X$  (soporte de  $X$ ).

Ahora, si  $X$  es variable aleatoria continua, las probabilidades están asociadas a intervalos de  $x$ .

En este caso se define la función de densidad  $f_X(x)$  tal que

$$P(a < X \leq b) = \int_a^b f_X(x) dx \quad (1)$$

y

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(t) dt$$

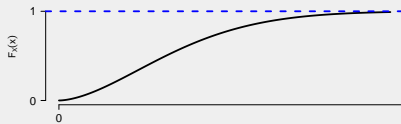
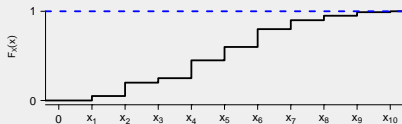
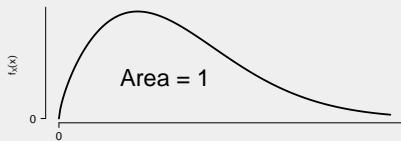
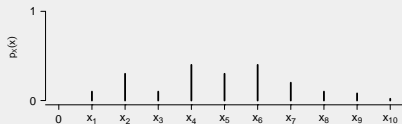
con

$$f_X(x) = \frac{d}{dx} F_X(x)$$

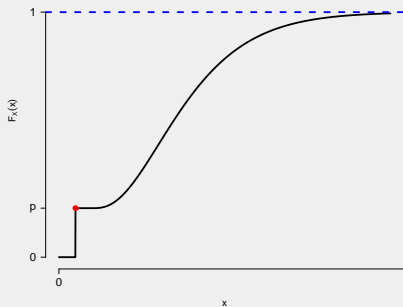
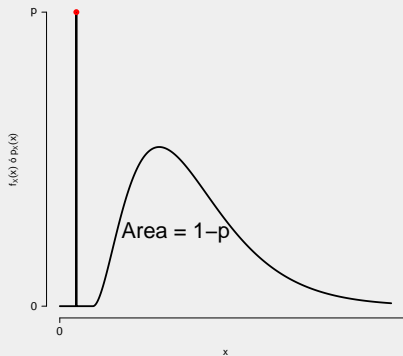
Notar que

$$P(x < X \leq x + dx) = f_X(x) dx$$

## Caso Discreto y Continuo



## Caso Mixto



## Propiedades

1.  $F_X(-\infty) = 0$  y  $F_X(\infty) = 1$ .
2.  $F_X(x) \geq 0$  para todo valor de  $x$  y es no decreciente.
3.  $F_X(x)$  es continua por la derecha

Para el caso continuo, la ecuación (1) la podemos escribir como

$$P(a < X \leq b) = \int_{-\infty}^b f_X(x) dx - \int_{-\infty}^a f_X(x) dx$$

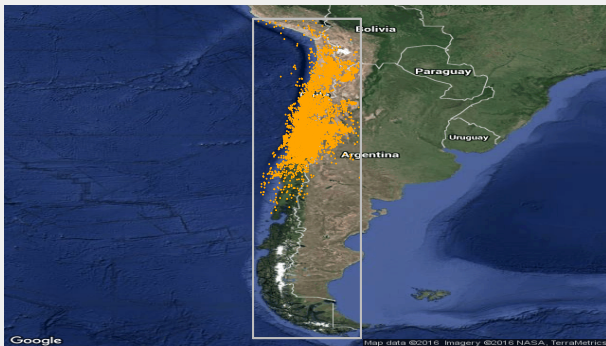
mientras que en el caso discreto

$$P(a < X \leq b) = \sum_{x_i \leq b} p_X(x_i) - \sum_{x_i \leq a} p_X(x_i)$$

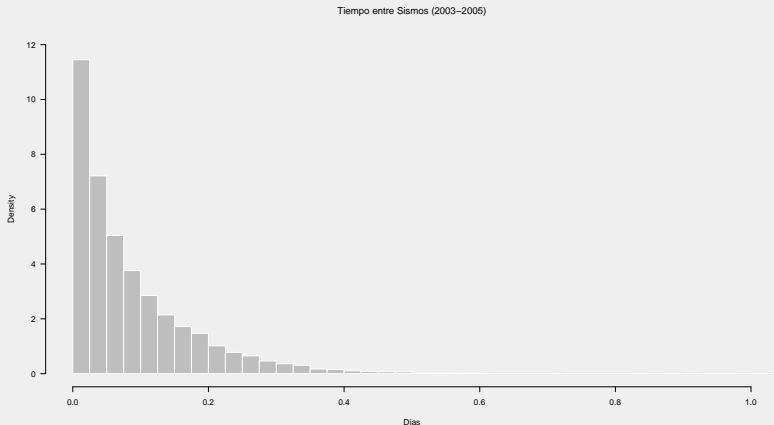
es decir, para ambos casos

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

**Aplicación:** Sismos ocurridos en Chile entre los años 2003 y 2005.



**Ejemplo 3.1** Definamos como  $X$  al tiempo entre sismos ocurridos en nuestro país.





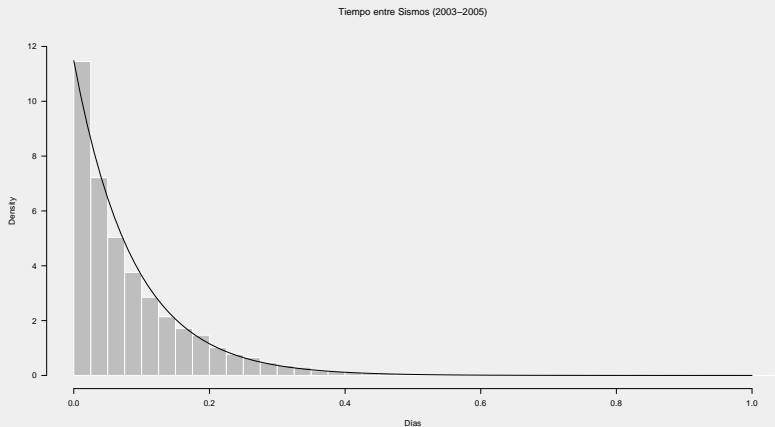
Este comportamiento puede ser descrito con una distribución cuya función de densidad es de la siguiente forma:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{en otro caso} \end{cases}$$

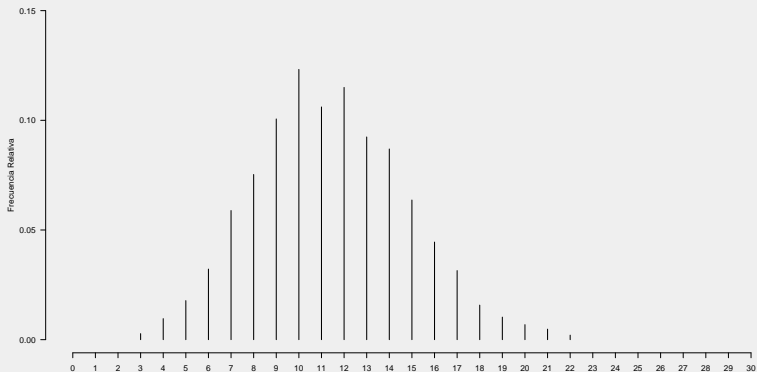
en que  $\lambda$  es una constante positiva. Esta distribución se conoce como distribución **Exponencial**.

Las funciones **dexp()** y **pexp()** de R, evalúan la función de densidad y de probabilidad acumulada para un valor en los reales.

# VARIABLES ALEATORIAS Y DISTRIBUCIÓN DE PROBABILIDAD



**Ejemplo 3.2** Definamos como  $Y$  al número de sismos diarios ocurridos en nuestro país.



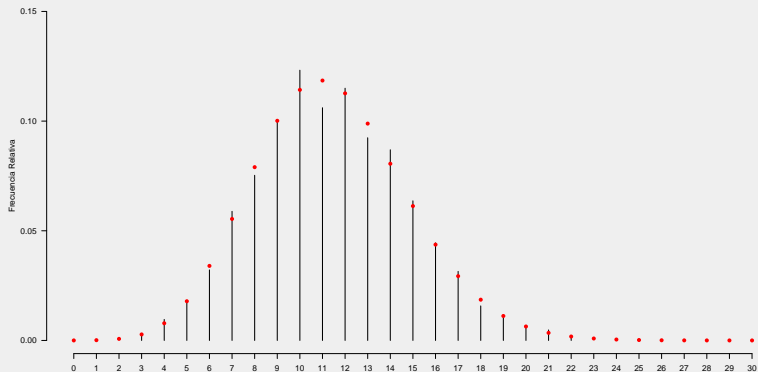
Este comportamiento puede ser descrito con una distribución cuya función de probabilidad es de la siguiente forma:

$$p_Y(y) = \begin{cases} \frac{\lambda^y e^{-\lambda}}{y!}, & y \in \mathbb{N}_0 \\ 0, & \text{en otro caso} \end{cases}$$

en que  $\lambda$  es una constante positiva. Esta distribución se conoce como distribución **Poisson**.

Las funciones **dpois()** y **ppois()** de R, evalúan la función de probabilidad y de probabilidad acumulada para un valor en los reales.

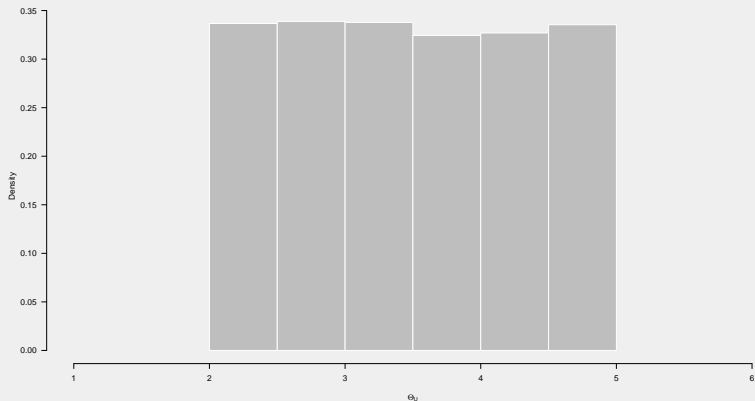
# VARIABLES ALEATORIAS Y DISTRIBUCIÓN DE PROBABILIDAD



**Ejemplo 3.3** Los números aleatorios son la base esencial de la simulación de variables aleatorias que veremos más adelante. Como primera experiencia supongamos que cada alumno de la universidad va a elegir totalmente al azar un número entre 2 y 5.

```
set.seed(1113) ## Semilla de simulación
U = runif(20000, 2, 5)
hist(U, freq = FALSE, col = "gray", border = "white",
las = 1, nclass = 5, main = "", xlab = expression(Theta[U]))
```

# VARIABLES ALEATORIAS Y DISTRIBUCIÓN DE PROBABILIDAD



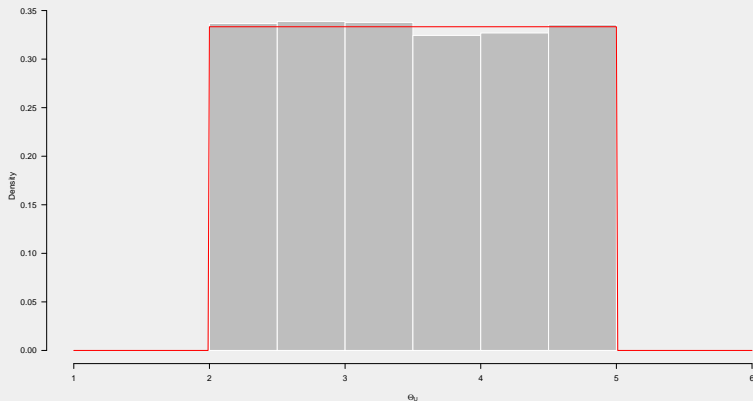
Una función densidad constante en el intervalo  $[2, 5]$  y cero en otro caso, ajusta de buena manera el comportamiento de los datos:

$$f_X(x) = \begin{cases} c, & 2 \leq x \leq 5 \\ 0, & \text{en otro caso} \end{cases}$$

Notar que la constante  $c$  para este caso es igual a  $\frac{1}{5-2}$ .



# VARIABLES ALEATORIAS Y DISTRIBUCIÓN DE PROBABILIDAD



A este modelo se le conoce como Uniforme( $a, b$ ), donde

$$f(u) = \frac{1}{b-a} \quad y \quad F(u) = \frac{u-a}{b-a}$$

para  $u \in [a, b]$

Una variable aleatoria puede ser descrita totalmente por su **función de distribución de probabilidad o de densidad**, o bien por su **función de distribución de probabilidad acumulada**.

Sin embargo, en la práctica la forma exacta puede no ser totalmente conocida.

En tales casos se requieren ciertas “medidas” para tener una idea de la forma de la distribución:

- Medidas Centrales.
- Medidas de Posición.
- Medidas de Dispersión.
- Medidas de Asimetrías.
- Otras

## Valores Centrales

En el rango de posibles valores de una variable aleatoria, existe un interés natural con respecto a los valores centrales, por ejemplo, el promedio.

Consideremos una variable aleatoria  $X$  con soporte  $\Theta_X$ .

Como cada valor de  $\Theta_X$  tiene una medida de probabilidad, el **promedio ponderado** es de especial interés.

## Valores Centrales

A el promedio ponderado se le llama también **valor medio** o **valor esperado** de la variable aleatoria  $X$ .

Para una variable aleatoria  $X$  se define el valor esperado,  $\mu_X$ , como:

$$\mu_X = E(X) = \begin{cases} \sum_{x \in \Theta_X} x \cdot p_X(x), & \text{Caso Discreto} \\ \int_{-\infty}^{\infty} x \cdot f_X(x) dx, & \text{Caso Continuo} \end{cases}$$

Este valor existe siempre y cuando

$$\sum_{x \in \Theta_X} |x| \cdot p_X(x) < \infty \quad \text{o} \quad \int_{-\infty}^{\infty} |x| \cdot f_X(x) dx < \infty$$

Otras medidas de centralidad corresponden:

- **La Moda:** Valor más frecuente o con mayor probabilidad.
- **La Mediana:** Sea  $x_{\text{med}}$  el valor que toma la mediana, entonces:

$$F_X(x_{\text{med}}) = 1/2$$

En resumen, el valor esperado de una variable aleatoria es un valor promedio que puede ser visto como un indicador del valor central de la distribución de probabilidad, por esta razón se considera como un parámetro de localización.

Por otra parte, la mediana y la moda de una distribución también son parámetros de localización que no necesariamente son iguales a la media.

Cuando la distribución es simétrica, estas tres medidas son parecidas.

## Medidas de Posición:

Entre las medidas de posición más usuales, se encuentra el mínimo, máximo y percentil  $p \times 100\%$ .

Si  $x_p$  es el valor que toma el percentil  $p \times 100\%$ , entonces  $F_X(x_p) = p$ .

Algunos casos particulares de percentil son: Quintiles, cuartiles, deciles, mediana.

## Esperanza matemática

La noción del valor esperado como un promedio ponderado puede ser generalizado para funciones de la variable aleatoria  $X$ .

Dada una función  $g(X)$ , entonces el valor esperado de esta puede ser obtenido como:

$$E[g(X)] = \begin{cases} \sum_{x \in \Theta_X} g(x) \cdot p_X(x), & \text{Caso Discreto} \\ \int_{-\infty}^{\infty} g(x) \cdot f_X(x) dx, & \text{Caso Continuo} \end{cases}$$

**Nota:** La demostración de este resultado se realizará durante el desarrollo del capítulo 4.



## **Función Generadora de Momentos** [Rice, 1995 (pág 142 - 144)]

La función generadora de momentos de una variable aleatoria  $X$  se define como

$$M_X(t) = E[\exp(tX)]$$

Esta función puede no estar definida para algunos valores de  $t$ , pero si existe en un intervalo abierto que contenga al cero, entonces esta función tiene la propiedad de determinar la distribución de probabilidad de  $X$ .

Cuando esto último ocurra, esta función permite obtener el  $r$ -ésimo momento de  $X$  de la siguiente forma

$$M^{(r)}(0) = E(X^r)$$

## Medidas de dispersión

Es de interés cuantificar el nivel de dispersión que tienen una variable aleatoria con respecto a un valor de referencia.

Por ejemplo, nos podría interesar la distancia esperada de los valores de una variable aleatoria  $X$  con respecto al valor esperado  $\mu_X$ , es decir,  $E[(X - \mu_X)]$ .

Esta idea de dispersión tiene el problema que siempre da como resultado cero.

## Medidas de dispersión

Una alternativa es utilizar la definición de varianza:

$$\begin{aligned}\sigma_X^2 &= \text{Var}(X) = E[(X - \mu_X)^2] \\ &= \begin{cases} \sum_{x \in \Theta_X} (x - \mu_X)^2 \cdot p_X(x), & \text{Caso Discreto} \\ \int_{-\infty}^{\infty} (x - \mu_X)^2 \cdot f_X(x) dx, & \text{Caso Continuo} \end{cases} \\ &= E(X^2) - \mu_X^2\end{aligned}$$

En términos de dimensionalidad, es conveniente utilizar la desviación estandar, es decir,

$$\sigma_X = \sqrt{\text{Var}(X)}$$

Ahora, si  $\mu_X > 0$ , una medida adimensional de la variabilidad es el coeficiente de variación (c.o.v)

$$\delta_X = \frac{\sigma_X}{\mu_X}$$

Otras medidas de dispersión son:

$$\text{Rango} = \max - \min$$

$$\text{IQR} = x_{0.75} - x_{0.25}$$

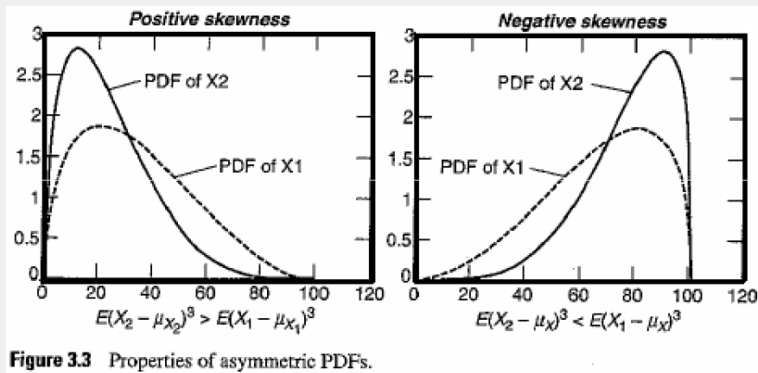
## Medida de asimetría

Se define una medida de asimetría (skewness) corresponde al tercer momento central:

$$E[(X - \mu_X)^3] = \begin{cases} \sum_{x_i \in \Theta_X} (x_i - \mu_X)^3 \cdot p_X(x_i), & \text{Caso Discreto} \\ \int_{-\infty}^{\infty} (x - \mu_X)^3 \cdot f_X(x) dx, & \text{Caso Continuo} \end{cases}$$

Una medida conveniente es el coeficiente de asimetría que se define como:

$$\theta_X = \frac{E[(X - \mu_X)^3]}{\sigma_X^3}$$



**Medida de kurtosis**

Finalmente, el cuarto momento central se conoce como la kurtosis

$$E[(X - \mu_X)^4] = \begin{cases} \sum_{x_i \in \Theta_X} (x_i - \mu_X)^4 \cdot p_X(x_i), & \text{Caso Discreto} \\ \int_{-\infty}^{\infty} (x - \mu_X)^4 \cdot f_X(x) dx, & \text{Caso Continuo} \end{cases}$$

que es una medida del “apuntamiento” o “achataamiento” de la distribución de probabilidad o de densidad.

Usualmente se prefiere el coeficiente de kurtosis

$$K_X = \frac{E[(X - \mu_X)^4]}{\sigma_X^4} - 3$$

**Ejercicio 2**

Para el modelo propuesto en el Ejemplo 3.1, muestre que:

(a)  $F_X(x) = 1 - e^{-\lambda x}, \quad x > 0.$

(b) Mediana =  $\frac{\ln(2)}{\lambda}.$

(c)  $\mu_X = \frac{1}{\lambda}.$

(d)  $M_X(t) = \frac{\lambda}{\lambda - t}, \quad t < \lambda.$

(e)  $\sigma_X^2 = \frac{1}{\lambda^2}.$

(f)  $\delta_X = 1.$



**Ejercicio 4** El tiempo  $T$  entre sismos vimos que se comportan como una variable aleatoria Exponencial (Ejemplo 3.1).

Si el tiempo medio es  $\mu_T$ , muestre que el tercer momento central es

$$E[(T - \mu_T)^3] = 2\mu_T^3$$

y que el coeficiente de asimetría es

$$\theta_T = 2$$

**Ejercicio 3**

Para el modelo propuesto en el Ejemplo 3.2, muestre que:

- (a)  $\mu_X = \lambda$ .
- (b)  $M_X(t) = \exp [\lambda (e^t - 1)]$ ,  $t \in \mathbb{R}$ .
- (c)  $\sigma_X^2 = \lambda$ .

**Ejemplo 3.4** Sean  $X$  e  $Y$  dos vectores de datos simulados:

```
set.seed(1113)
X = rgamma(10000, rate = 1/2, shape = 1/2)
Y = rgamma(10000, rate = 1, shape = 1)
par(mfrow = c(1,2), cex = 0.8)
hist(X, freq = FALSE, ylim = c(0,0.9), xlim = c(0,10),
breaks = seq(0,25,1), main = "Histograma de X")
hist(Y, freq = FALSE, ylim = c(0,0.9), xlim = c(0,10),
breaks = seq(0,25,1), main = "Histograma de Y")
```

¿Cuál de los dos se comporta como una variable aleatoria Exponencial?

```
Tabla = rbind(  
  c(mean(X), sd(X), sd(X)/mean(X), mean(((X-mean(X))/sd(X))^3)),  
  c(mean(Y), sd(Y), sd(Y)/mean(Y), mean(((Y-mean(Y))/sd(Y))^3))  
)  
colnames(Tabla) = c("mu", "sigma", "delta", "theta")  
rownames(Tabla) = c("X", "Y")  
round(Tabla,1)
```

# DISTRIBUCIONES DE PROBABILIDAD

Teóricamente, cualquier función que satisfaga las condiciones descritas puede ser utilizada para representar una distribución de probabilidad o de densidad de una variable aleatoria, sin embargo, nos interesan algunas funciones que satisfacen:

- La función es resultado de un proceso físico o puede ser derivada bajo ciertos supuestos.
- La función es resultado de algunos procesos límites
- Es ampliamente conocida, y la información esta disponible.
- Distribución de probabilidad, medidas de descripción, etc.

En estadística y probabilidad se llama distribución normal, distribución de Gauss o distribución gaussiana, a una de las distribuciones de probabilidad de variable continua que con más frecuencia aparece en fenómenos reales.

La distribución normal fue reconocida por primera vez por el francés **Abraham de Moivre** (1667-1754), en un artículo de 1733 donde relaciona la distribución normal y la binomial.

Posteriormente, el matemático alemán **Carl Friedrich Gauss** (1777-1855) elaboró desarrollos más profundos y formuló la ecuación de la curva; de ahí que asimismo se la conozca, más comúnmente, como campana de Gauss.

# DISTRIBUCIONES DE PROBABILIDAD





Muchos fenómenos que encontramos en la naturaleza se miden mediante variables cuya distribución, naturalmente, tienden a formar una campana de Gauss. Además, muchas variables aleatorias continuas presentan una función de densidad cuya gráfica tiene dicha forma.

Finalmente, varios procedimientos estadísticos usados habitualmente asumen la normalidad de los datos observados.

La importancia de esta distribución radica en que permite modelar numerosos fenómenos naturales, sociales y psicológicos.

La función densidad de una variable aleatoria  $X$  con distribución Normal( $\mu, \sigma$ ) es de la forma:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2 \right\}, \quad -\infty < x < \infty$$

con  $\mu$  parámetro de localización y  $\sigma$  un parámetro de escala o forma tales que:

$$-\infty < \mu < \infty, \quad 0 < \sigma < \infty$$

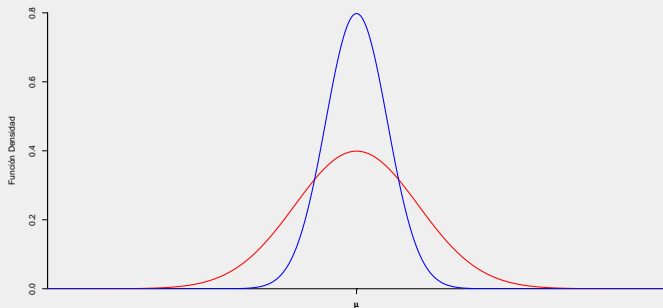
# DISTRIBUCIONES DE PROBABILIDAD

Los dos parámetros:  $\mu$  y  $\sigma$ , permiten describir diferentes distribuciones normales.

**Efecto del parámetro  $\mu$**



## Efecto del parámetro $\sigma$



Un caso especial es cuando  $\mu = 0$  y  $\sigma = 1$ . Este caso es conocido como la distribución normal estándar.

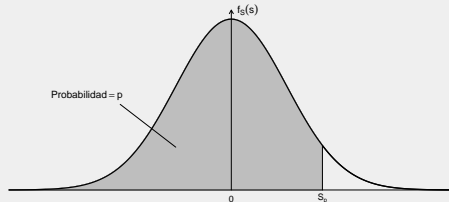
$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

La ventaja es que función de distribución de probabilidad acumulada se encuentra tabulada, la cual se denota por  $\Phi(\cdot)$ .

# DISTRIBUCIONES DE PROBABILIDAD

Sea  $S$  una variable aleatoria con distribución normal estándar, cuya función de distribución de probabilidad acumulada esta dada por

$$\Phi(s) = F_S(s) = \int_{-\infty}^s \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$



**Propiedades:**  $S_p = \Phi^{-1}(p) = -\Phi^{-1}(1 - p)$  y  $\Phi(-s) = 1 - \Phi(s)$ .

# DISTRIBUCIONES DE PROBABILIDAD

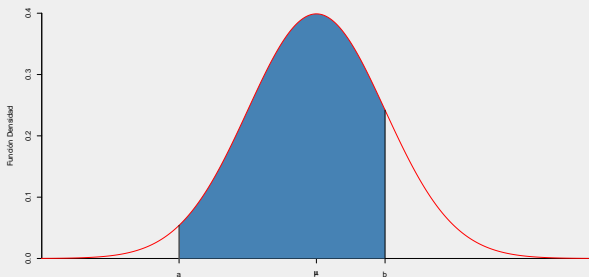
**Tabla Normal Estándar:** Resultado de  $\Phi(S_p) = p$ , para  $S_p \geq 0$

$S_p$	Segunda cifra decimal									
	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	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767

# DISTRIBUCIONES DE PROBABILIDAD

Sea  $X$  una variable aleatoria Normal( $\mu, \sigma$ ) con función de distribución acumulada  $F_X$ . Para dos valores dados  $a$  y  $b$  (con  $a < b$ ) se tiene que:

$$P(a < X \leq b) = F_X(b) - F_X(a)$$





**Ejercicio 5** Si  $X \sim \text{Normal}(\mu, \sigma)$ , muestre que:

1.  $E(X) = \mu$ .
2.  $\text{Var}(X) = \sigma^2$ .
3.  $F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ .

La función `dnorm()`, `qnorm()` y `pnorm()` de R, evalúan la densidad, calculan percentiles y probabilidades acumuladas respectivamente.

**Ejercicio 6** Durante una tormenta el drenaje de una comunidad se comporta como una variable aleatoria Normal con una media y desviación estándar estimada de 1.2 y 0.4 millones de galones diarios (mgd) respectivamente.

- (a) Si el sistema de drenaje de aguas pluviales fue diseñado para soportar una capacidad máxima de 1.5 mgd. ¿Cuál es la probabilidad que el sistema colapse?
- (b) ¿Cuál es la probabilidad que el sistema deba soportar un drenaje entre 1.0 y 1.6 mgd?
- (c) ¿Cuál es la carga de agua durante una tormenta que soporta el sistema de drenaje correspondiente al percentil 90?

Se dice que  $X$  sigue una ley de probabilidad Log-Normal si su función de densidad esta dada por

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{(\zeta x)} \exp \left[ -\frac{1}{2} \left( \frac{\ln x - \lambda}{\zeta} \right)^2 \right], \quad x \geq 0.$$

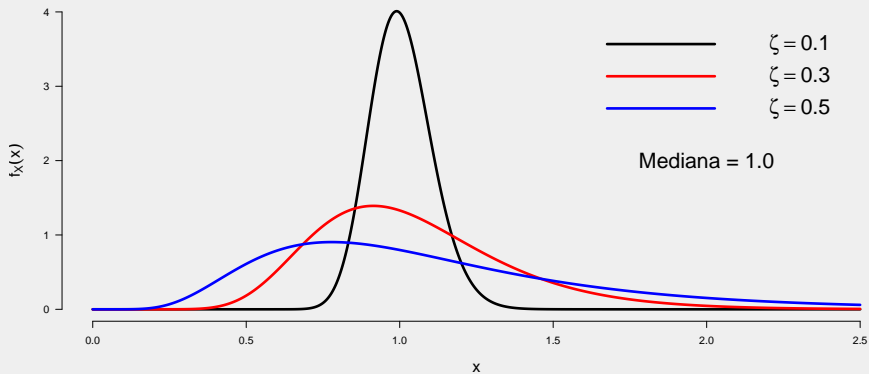
Donde,

$$\lambda = E(\ln X) \quad \text{y} \quad \zeta = \sqrt{\text{Var}(\ln X)}$$

**Propiedad:**  $\ln X \sim \text{Normal}(\lambda, \zeta)$  (Demostración en el capítulo 4)

La función *dlnorm()*, *qlnorm()* y *plnorm()* de R, evalúan la densidad, calculan percentiles y probabilidades acumuladas respectivamente en el caso Log-Normal.

# DISTRIBUCIONES DE PROBABILIDAD

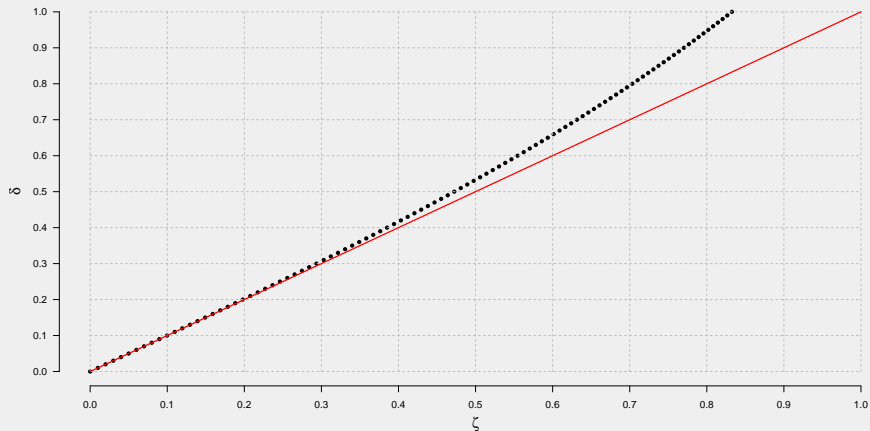


## Ejercicio 7

Sea  $X \sim \text{Log-Normal}(\lambda, \zeta)$ , muestre que

- $\mu_X = \exp(\lambda + \zeta^2/2)$ .
- Mediana =  $\exp(\lambda)$
- $E(X^k) = \exp(\lambda k) \cdot M_Z(\zeta k)$ , con  $Z \sim \text{Normal}(0,1)$ .
- $\sigma_X^2 = \mu_X^2(e^{\zeta^2} - 1)$ .
- $\zeta = \sqrt{\ln(1 + \delta_X^2)}$

Relación  $\delta$  (c.o.v) vs.  $\zeta$



**Ejercicio 8** Considere ahora que la distribución del drenaje se comporta como una variable aleatoria Log-Normal en vez de Normal.

- (a) Si el sistema de drenaje de aguas pluviales fue diseñado para soportar una capacidad máxima de 1.5 mgd. ¿Cuál es la probabilidad que el sistema colapse?
- (b) ¿Cuál es la probabilidad que el sistema deba soportar un drenaje entre 1.0 y 1.6 mgd?
- (c) ¿Cuál es la carga de agua durante una tormenta que soporta el sistema de drenaje correspondiente al percentil 90?

En las más diversas áreas de la Ingeniería, a menudo los problemas involucran la ocurrencia o recurrencia de un evento, el cual es impredecible, como una secuencia de “experimentos”.

## **Ejemplos:**

1. Para un día de lluvia, ¿colapsa o no un sistema de drenaje?
2. Al comprar un producto, ¿éste satisface o no los requerimientos de calidad?
3. Un alumno ¿aprueba o reprueba el curso?

Notar que hay sólo dos resultados posibles para cada “experimento”.



Las variables descritas pueden ser modeladas por una secuencia Bernoulli, la cual se basa en los siguientes supuestos:

1. Cada experimento, tiene una de dos opciones ocurrencia o no ocurrencia del evento.
2. La probabilidad de ocurrencia del evento (“éxito”) en cada experimento es constante (digamos  $p$ ).
3. Los experimentos son estadísticamente independientes.

# DISTRIBUCIONES DE PROBABILIDAD

Dada una secuencia Bernoulli, si  $X$  es el número de ocurrencias del evento éxito entre los  $n$  experimentos, con probabilidad de ocurrencia igual a  $p$ , entonces la probabilidad que ocurran exactamente  $x$  éxitos en los  $n$  experimentos esta representada por la distribución Binomial

$$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \dots, n$$

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \sum_{k=0}^{[x]} \binom{n}{k} p^k (1-p)^{n-k}, & 0 \leq x < n \\ 1, & x \geq n \end{cases}$$

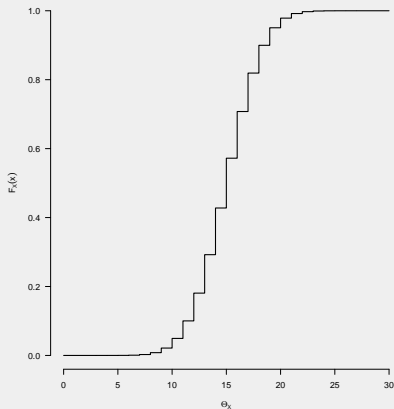
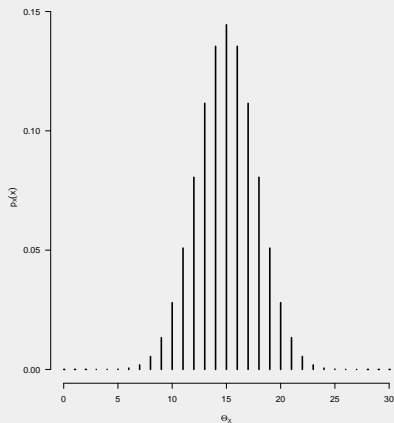
El valor esperado y varianza están dados por:

$$E(X) = np, \quad \text{Var}(X) = np(1 - p)$$

Las funciones *dbinom()*, *qbinom()* y *pbinom()* de R, evalúan la función de probabilidad, calcula percentiles y probabilidades acumuladas respectivamente en el caso Binomial( $n, p$ ).

# DISTRIBUCIONES DE PROBABILIDAD

Binomial( $n = 30, p = 1/2$ )



**Ejercicio 9** Cinco niveladoras se utilizan en la construcción de una carretera. Suponga que  $T$ , vida operacional de cada niveladora, se comporta como una variable aleatoria log-normal con media 1.500 hrs. y c.o.v. del 30%.

Asumiendo independencia entre las niveladoras, determine la probabilidad que:

- (a) exactamente dos presenten una falla en las primeras 900 hrs. de funcionamiento.
- (b) dos o más presenten fallas en las primeras 900 hrs.
- (c) a lo más dos presenten fallas en las primeras 900 hrs.

Dada una secuencia Bernoulli, el número de experimentos hasta la ocurrencia del primer evento exitoso sigue una distribución geométrica.

Si el primer éxito ocurre en el  $n$ -ésimo experimento, los primeros  $n - 1$  fueron “fracasos”. Si  $N$  es la variable aleatoria que representa el número de experimentos hasta el primer éxito, entonces:

$$P(N = n) = p(1 - p)^{n-1}, \quad n = 1, 2, \dots$$

La función distribución esta dada por:

$$F_N(n) = \sum_{k=1}^{[n]} p(1-p)^{k-1} = 1 - (1-p)^{[n]}$$

para  $n > 0$ .

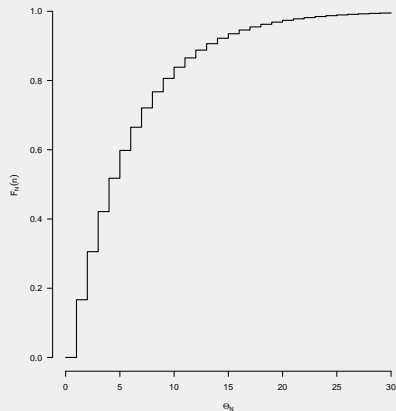
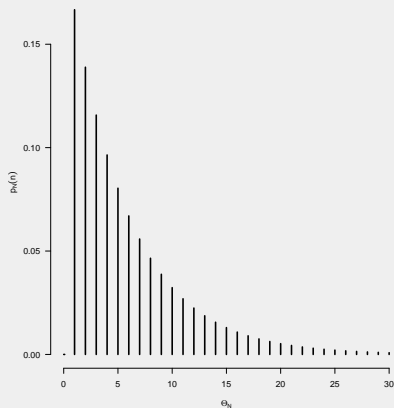
Mientras que su valor esperado y varianza son:

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

Las funciones *dgeom()*, *qgeom()* y *pgeom()* de R, evalúan la función de probabilidad, calcula percentiles y probabilidades acumuladas respectivamente en el caso Geométrica( $p$ ).

# DISTRIBUCIONES DE PROBABILIDAD

Geométrica( $p = 1/6$ )





## Tiempo de recurrencia y periodo de retorno

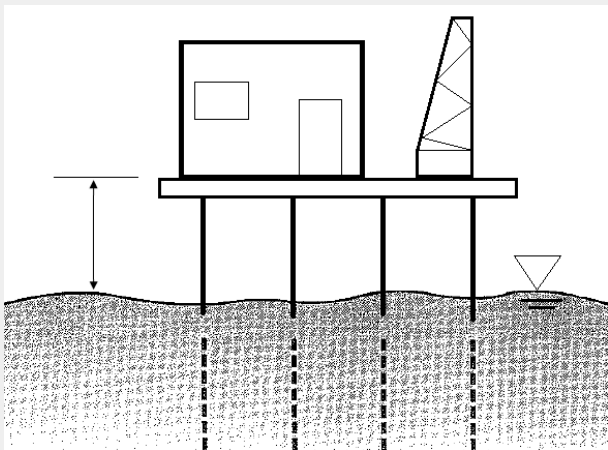
En algunos problemas, el tiempo  $T$  (o espacio) es discretizado en intervalos,  $T = N$ , y puede ser modelado por secuencia Bernoulli.

El número de intervalos ocurridos hasta observar el primer evento exitoso se denomina tiempo medio de recurrencia y corresponde al tiempo entre recurrencias. Este tiempo se conoce como periodo de retorno:

$$\bar{T} = E(T) = \sum_{t=1}^{\infty} t p (1-p)^{t-1} = \frac{1}{p}$$

# DISTRIBUCIONES DE PROBABILIDAD

**Ejercicio 10** Una plataforma “costa afuera” es diseñada para soportar olas de hasta 8 metros sobre el nivel medio del mar (ver figura).



Una ola de estas característica tiene una probabilidad del 5% de ocurrencia por año.

Determine:

- El periodo de retorno.
- La probabilidad de observar una ola de estas características durante el período de retorno.
- Probabilidad que ocurra después del tercer año.
- Probabilidad que ocurra el 5º año, dado que no sucedió durante los primeros tres años.

La distribución geométrica permite modelar el numero de experimentos hasta la primera ocurrencia.

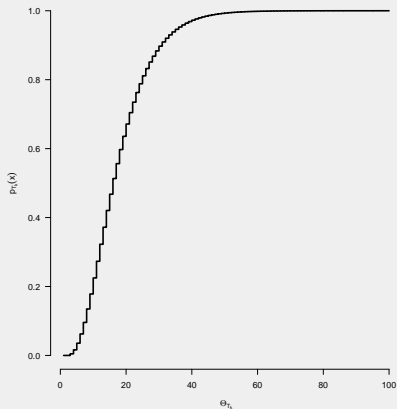
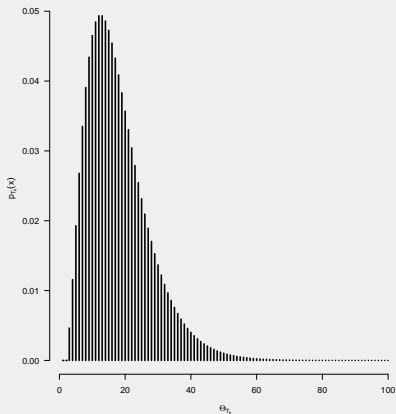
El numero de experimentos hasta la  $k$ -ésima ocurrencia de un éxito es modelada por la distribución binomial negativa.

$$P(T_k = x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}, \quad x = k, k+1, k+2, \dots$$

$$E(T_k) = \frac{k}{p}, \quad \text{Var}(T_k) = \frac{k(1-p)}{p^2}$$

# DISTRIBUCIONES DE PROBABILIDAD

Bin-Neg( $k = 3, p = 1/6$ )



Muchos problemas físicos de interés para ingenieros y científicos que implican las ocurrencias posibles de eventos en cualquier punto en el tiempo y/o en el espacio.

**Por ejemplo:**

- Los terremotos pueden ocurrir en cualquier momento y en cualquier lugar en una región con actividad sísmica en el mundo.
- Las grietas por fatiga puede producirse en cualquier punto de una soldadura continua.
- Los accidentes de tráfico pueden suceder en cualquier momento en una autopista.

Este problema puede ser modelado como secuencia Bernoulli, dividiendo el tiempo o el espacio en pequeños intervalos “apropiados” tal que solo un evento puede ocurrir o no dentro de cada intervalo (Ensayo Bernoulli).

Sin embargo, si el evento puede ocurrir al azar en cualquier instante de tiempo (o en cualquier punto del espacio), esto puede ocurrir más de una vez en cualquier momento o intervalo de espacio.

En tal caso, las ocurrencias del evento puede ser más apropiado el modelo con un proceso de Poisson o la secuencia Poisson.

## Supuestos:

- Un evento puede ocurrir al azar y en cualquier instante de tiempo o en cualquier punto en el espacio.
- La ocurrencia(s) de un evento en un intervalo de tiempo dado (o espacio) es estadísticamente independiente a lo que ocurra en otros intervalos (o espacios) que no se solapen.
- La probabilidad de ocurrencia de un evento en un pequeño intervalo  $\Delta t$  es proporcional a  $\Delta t$ , y puede estar dada por  $\nu \Delta t$ , donde  $\nu$  es la tasa de incidencia media del evento (que se supone constante).
- La probabilidad de dos o más eventos en  $\Delta t$  es insignificante.



Bajo los supuestos anteriores, el número de eventos estadísticamente independientes en  $t$  (tiempo o espacio) está regido por la función de probabilidad del modelo Poisson, donde la variable aleatoria  $X_t$ : número de eventos en el intervalo de tiempo  $(0, t)$ .

$$P(X_t = x) = \frac{(\nu t)^x e^{-\nu t}}{x!} = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \dots$$

donde  $\nu$  es la tasa de ocurrencia media por unidad de tiempo y  $\lambda$  su esperanza en  $(0, t)$ :

$$E(X_t) = \nu t = \lambda$$

En R, la función `dpois()` evalúa la función de probabilidad y las funciones: `qpois()` y `ppois()`, entregan percentiles y probabilidades acumuladas respectivamente.

**Ejercicio 11** Los registros históricos de tormentas severas en una ciudad en los últimos 20 años indica que ha habido un promedio de cuatro tormentas de lluvia por año.

Suponiendo que las ocurrencias de lluvias se pueden modelar con un Proceso de Poisson.

Calcule la probabilidad el próximo año no presente tormentas de lluvias.

En un Proceso de Poisson el tiempo transcurrido entre la ocurrencia de eventos puede ser descrito por una distribución exponencial.

Si  $T_1$  representa al tiempo transcurrido hasta la ocurrencia del primer evento en un Proceso de Poisson, el evento  $(T_1 > t)$  implica que en el intervalo  $(0, t)$  no ocurren eventos, es decir,

$$P(T_1 > t) = P(X_t = 0) = \frac{(\nu t)^0 e^{-\nu t}}{0!} = e^{-\nu t},$$

con

$$X_t \sim \text{Poisson}(\nu t)$$

Por lo tanto la función de distribución de probabilidad acumulada de  $T_1$  esta dada por:

$$F_{T_1}(t) = P(T_1 \leq t) = 1 - P(T_1 > t) = 1 - e^{-\nu t}$$

Su función densidad se obtiene como sigue:

$$f_{T_1}(t) = \frac{d}{dt}F_{T_1}(t) = \nu e^{-\nu t}$$

que corresponde a la función densidad de una variable aleatoria con distribución exponencial.

Esta distribución al igual que la geométrica tiene la propiedad de la carencia de memoria, es decir, si  $T \sim \text{Exponencial}(\nu)$  entonces se tiene que

$$P(T > t + s \mid T > s) = P(T > t)$$

Este resultado, nos permite asumir que todos los tiempo entre eventos  $\text{Poisson}(\nu t)$  distribuyen  $\text{Exponencial}(\nu)$ .

Ya mostramos anteriormente que una variable aleatoria  $X$  con distribución Exponencial de parámetro  $\nu > 0$ , tiene función densidad y de distribución:

$$f_X(x) = \begin{cases} \nu e^{-\nu x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad F_X(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\nu x}, & x \geq 0 \end{cases}$$

Mientras que su valor esperado y varianza son:

$$\mu_X = \frac{1}{\nu}, \quad \sigma_X^2 = \frac{1}{\nu^2}$$

En R, la función `dexp()` evalúa la función de densidad y las funciones: `qexp()` y `pexp()`, entregan percentiles y probabilidades acumuladas respectivamente.

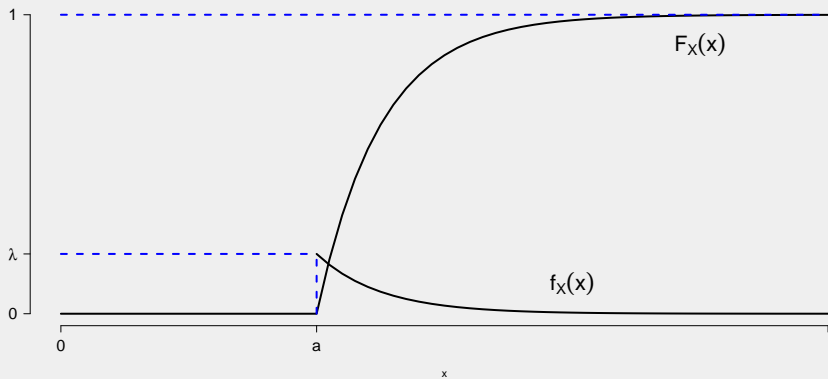
Una variable aleatoria  $X$  con distribución Exponencial de parámetro  $\nu > 0$ , se llama trasladada en  $a$  si su función densidad y de distribución acumulada son

$$f_X(x) = \begin{cases} \nu e^{-\nu(x-a)}, & x \geq a \\ 0, & x < a \end{cases} \quad F_X(x) = \begin{cases} 0, & x < a \\ 1 - e^{-\nu(x-a)}, & x \geq a \end{cases}$$

Su valor esperado y varianza están dados por

$$\mu_X = \frac{1}{\nu} + a, \quad \sigma_X^2 = \frac{1}{\nu^2}$$

# DISTRIBUCIONES DE PROBABILIDAD





Una variable aleatoria  $X$  con distribución Gamma tiene función densidad

$$f_X(x) = \frac{\nu^k}{\Gamma(k)} x^{k-1} e^{-\nu x}, \quad x \geq 0$$

donde  $k, \nu$  son parámetros positivos.

La función  $\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u} du$ , la cual tiene las siguientes propiedades:

- $\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$ .
- $\Gamma(n + 1) = n!$  si  $n \in \mathbb{N}_0$ .
- $\Gamma(1/2) = \sqrt{\pi}$ .

En R, la función `dgamma()` evalúa la función de densidad y las funciones: `qgamma()` y `pgamma()`, entregan percentiles y probabilidades acumuladas respectivamente.

## Relación con distribución Poisson

En un Proceso de Poisson el tiempo transcurrido hasta la ocurrencia del  $k$ -ésimo evento puede ser descrito por una distribución Gamma.

Si  $T_k$  representa al tiempo transcurrido hasta la ocurrencia del  $k$ -ésimo evento en un Proceso de Poisson, el evento  $(T_k > t)$  implica que en el intervalo  $(0, t)$  ocurren a lo más  $k-1$  eventos, es decir,

$$P(T_k > t) = P(X_t \leq k - 1) = \sum_{x=0}^{k-1} \frac{(\nu t)^x e^{-\nu t}}{x!}$$

## Relación con distribución Poisson

Luego, su función de distribución acumulada esta dada por:

$$F_{T_k}(t) = 1 - \sum_{x=0}^{k-1} \frac{(\nu t)^x e^{-\nu t}}{x!}$$

Se puede demostrar que

$$f_{T_k}(t) = \frac{d}{dt} F_{T_k}(t) = \frac{\nu^k}{\Gamma(k)} t^{k-1} e^{-\nu t}, \quad t \geq 0$$

donde su valor esperado y varianza son

$$\mu_{T_k} = \frac{k}{\nu}, \quad \sigma_{T_k}^2 = \frac{k}{\nu^2}$$

**Ejemplo 3.5** Supongamos que los accidentes mortales en una carretera particular se producen en promedio una vez cada 6 meses.

Si podemos suponer que la incidencia de los accidentes en esta carretera constituye un proceso de Poisson, con tasa de incidencia media de  $1/6$  accidentes por mes, el tiempo  $T_1$  hasta la aparición del primer accidente puede describirse como una variable aleatoria exponencial con función densidad:

$$f_{T_1}(t) = \frac{1}{6} e^{-t/6}$$

El tiempo  $T_2$  hasta la ocurrencia del segundo accidente puede describirse como una variable aleatoria Gamma con función de densidad:

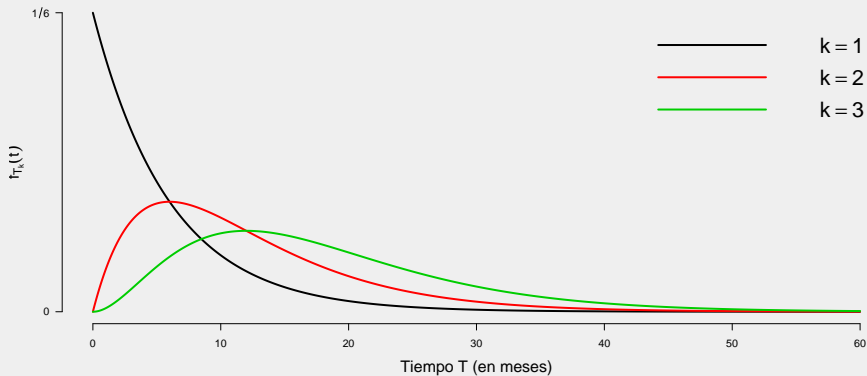
$$f_{T_2}(t) = \frac{(1/6)^2}{\Gamma(2)} t^{2-1} e^{-t/6}$$

Y el tiempo  $T_3$  hasta la ocurrencia del tercer accidente como una variable aleatoria Gamma con función de densidad:

$$f_{T_3}(t) = \frac{(1/6)^3}{\Gamma(3)} t^{3-1} e^{-t/6}$$

# DISTRIBUCIONES DE PROBABILIDAD

Distribución Gamma( $k, \nu = 1/6$ )



Una variable aleatoria  $X$  tiene distribución Gamma trasladada si su función de densidad esta dada por

$$f_X(x) = \frac{\nu^k}{\Gamma(k)} (x - \gamma)^{k-1} e^{-\nu(x-\gamma)}, \quad x \geq \gamma$$

donde  $k$ ,  $\mu$  y  $\gamma$  son parámetros de la distribución.

Su valor esperado y varianza son:

$$\mu_X = \frac{k}{\nu} + \gamma, \quad \sigma_X^2 = \frac{k}{\nu^2}$$

# DISTRIBUCIONES DE PROBABILIDAD

La distribución **Hipergeométrica** se produce cuando se estudia una muestra de una población finita la cual se puede estar dividida en dos grupos.

Considere un lote de tamaño  $N$ , de los  $m$  son defectuosos y  $N - m$  no defectuosos.

Si se toma una muestra aleatoria de tamaño  $n$  al azar, la probabilidad que  $x$  sean defectuosos está dada por la función de probabilidad de la distribución hipergeométrica:

$$p_X(x) = \frac{\binom{m}{x} \binom{N-m}{n-x}}{\binom{N}{n}}, \quad \max\{0, n+m-N\} \leq x \leq \min\{n, m\}$$

con

$$X \sim \text{Hipergeométrica}(n, N, m)$$



El cálculo de su valor esperado y varianza requiere un desarrollo bastante complejo cuyo resultado final es el siguiente

$$\mu_X = n \cdot \frac{m}{N}, \quad \sigma_X^2 = \left( \frac{N-n}{N-1} \right) \cdot n \cdot \frac{m}{N} \cdot \left( 1 - \frac{m}{N} \right)$$

En R, la función *dhyper()* evalúa la función de probabilidad y las funciones: *qhyper()* y *phyper()*, entregan percentiles y probabilidades acumuladas respectivamente.

**Ejemplo 3.6** En una caja de 100 medidores de tensión, se sospecha que puede haber cuatro que están defectuosos.

Si seis de los medidores de la caja se utilizaron en un experimento, la probabilidad de que un medidor defectuoso fuese usado se evalúa como siguiente.

Tenemos que

$$N = 100, \quad m = 4 \quad n = 6$$

Luego

$$P(X = 1) = \frac{\binom{4}{1} \binom{100 - 4}{6 - 1}}{\binom{100}{6}} = 0.205$$

# DISTRIBUCIONES DE PROBABILIDAD

Una variable aleatoria  $X$  con distribución Beta tiene función densidad

$$f_X(x) = \frac{1}{B(r, q)} \cdot \frac{(x - a)^{q-1} (b - x)^{r-1}}{(b - a)^{q+r-1}}, \quad a \leq x \leq b$$

donde  $q$  y  $r$  son los parámetros de la distribución, y  $B(q, r)$  es la función beta dada por

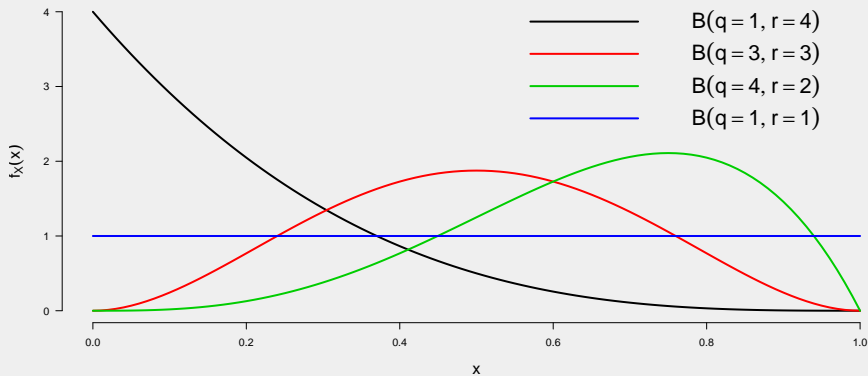
$$B(q, r) = \int_0^1 x^{q-1} (1 - x)^{r-1} dx = \frac{\Gamma(q) \Gamma(r)}{\Gamma(q + r)}$$

$$\mu_X = a + \frac{q}{(q + r)} (b - a) \quad \sigma_X^2 = \frac{q r (b - a)^2}{(q + r)^2 (q + r + 1)}$$

En R, la función `dbeta()` evalúa la función de densidad y las funciones: `qbeta()` y `pbeta()`, entregan percentiles y probabilidades acumuladas respectivamente.

# DISTRIBUCIONES DE PROBABILIDAD

Distribución Beta( $q, r$ )



# OTRAS DISTRIBUCIONES DE PROBABILIDAD

Si  $T \sim \text{Weibull}(\eta, \beta)$ , se tiene que

$$F_T(t) = 1 - \exp \left[ - \left( \frac{t}{\eta} \right)^\beta \right]$$
$$f_T(t) = \frac{\beta}{\eta} \left( \frac{t}{\eta} \right)^{\beta-1} \exp \left[ - \left( \frac{t}{\eta} \right)^\beta \right], \quad t > 0$$

Con  $\beta > 0$ , es un parámetro de forma y  $\eta > 0$ , es un parámetro de escala.

Si  $t_p$  es el percentil  $p \times 100\%$ , entonces

$$\ln(t_p) = \ln(\eta) + \frac{1}{\beta} \cdot \Phi_{\text{Weibull}}^{-1}(p), \quad \Phi_{\text{Weibull}}^{-1}(p) = \ln[-\ln(1 - p)]$$

Mientras que su  $m$ -ésimo momento está dado por

$$E(T^m) = \eta^m \Gamma(1 + m/\beta)$$

Luego

$$\mu_T = \eta \Gamma\left(1 + \frac{1}{\beta}\right), \quad \sigma_T^2 = \eta^2 \left[ \Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right) \right]$$

En R, la función *dweibull()* evalúa la función de densidad y las funciones: *qweibull()* y *pweibul()*, entregan percentiles y probabilidades acumuladas respectivamente.

Si  $Y \sim \text{Logística}(\mu, \sigma)$ , se tiene que

$$F_Y(y) = \Phi_{\text{Logística}}\left(\frac{y - \mu}{\sigma}\right); \quad f_Y(y) = \frac{1}{\sigma} \phi_{\text{Logística}}\left(\frac{y - \mu}{\sigma}\right), \quad -\infty < y < \infty$$

donde

$$\Phi_{\text{Logística}}(z) = \frac{\exp(z)}{[1 + \exp(z)]} \quad \text{y} \quad \phi_{\text{Logística}}(z) = \frac{\exp(z)}{[1 + \exp(z)]^2}$$

son la función de probabilidad y de densidad de una Logística Estándar.  $\mu \in \mathbb{R}$ , es un parámetro de localización y  $\sigma > 0$ , es un parámetro de escala.



Si  $y_p$  es el percentil  $p \times 100\%$ , entonces

$$y_p = \mu + \sigma \Phi_{\text{Logística}}^{-1}(p) \quad \text{con} \quad \Phi_{\text{Logística}}^{-1}(p) = \log \left( \frac{p}{1-p} \right)$$

Su esperanza y varianza están dadas por:  $\mu_Y = \mu$  y  $\sigma_Y^2 = \frac{\sigma^2 \pi^2}{3}$ .

En R, la función *dlogis()* evalúa la función de densidad y las funciones: *qlogis()* y *plogis()*, entregan percentiles y probabilidades acumuladas respectivamente.

Si  $T \sim \text{Log-Logística}(\mu, \sigma)$ , se tiene que

$$F_T(t) = \Phi_{\text{Logística}}\left(\frac{\ln(t) - \mu}{\sigma}\right); \quad f_T(t) = \frac{1}{\sigma t} \phi_{\text{Logística}}\left(\frac{\ln(t) - \mu}{\sigma}\right) \quad t > 0$$

$\exp(\mu)$ , es un parámetro de escala y  $\sigma > 0$ , es un parámetro de forma.

Si  $t_p$  es el percentil  $p \times 100\%$ , entonces

$$\ln(t_p) = \mu + \sigma \Phi_{\text{Logística}}^{-1}(p)$$

## Momentos:

Para un entero  $m > 0$  se tiene que

$$E(T^m) = \exp(m\mu) \Gamma(1 + m\sigma) \Gamma(1 - m\sigma)$$

El  $m$ -ésimo momento no es finito si  $m\sigma \geq 1$ .

Para  $\sigma < 1$

$$\mu_T = \exp(\mu) \Gamma(1 + \sigma) \Gamma(1 - \sigma)$$

y para  $\sigma < 1/2$

$$\sigma_T^2 = \exp(2\mu) [\Gamma(1 + 2\sigma) \Gamma(1 - 2\sigma) - \Gamma^2(1 + \sigma) \Gamma^2(1 - \sigma)]$$

Un variable aleatoria  $T$  tiene distribución t-student si su función de densidad está dada por:

$$f_T(t) = \frac{\Gamma[(\nu + 1)/2]}{\sqrt{\pi \nu} \Gamma(\nu/2)} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}, \quad -\infty < t < \infty$$

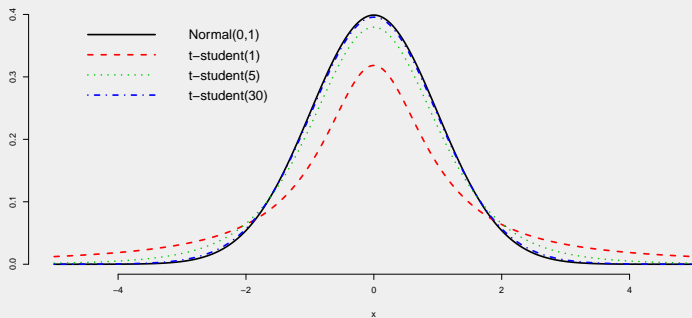
W. S. Gosset (1908).

- $\mu_T = 0$ , para  $\nu > 1$ .
- $\sigma_T^2 = \frac{\nu}{\nu-2}$ , para  $\mu > 2$ .

En R, la función  $dt()$  evalúa la función de densidad y las funciones:  $qt()$  y  $pt()$ , entregan percentiles y probabilidades acumuladas respectivamente.

# OTRAS DISTRIBUCIONES DE PROBABILIDAD

## Normal Estándar vs. t-Student



Si  $T \sim \text{Fisher}(\eta, \nu)$ , se tiene que

$$f_T(t) = \frac{\Gamma(\frac{\eta+\nu}{2})}{\Gamma(\eta/2)\Gamma(\nu/2)} \left(\frac{\eta}{\nu}\right)^{\frac{\eta}{2}} \frac{t^{\frac{\eta}{2}-1}}{\left(\frac{\eta}{\nu}t + 1\right)^{\frac{\eta+\nu}{2}}}, \quad t > 0$$

- $\mu_T = \frac{\nu}{\nu-2}$ , para  $\nu > 2$ .
- $\sigma_T^2 = \frac{2\nu^2(\eta+\nu-2)}{\eta(\nu-2)^2(\nu-4)}$ , para  $\nu > 4$ .

En R, la función `df()` evalúa la función de densidad y las funciones: `qf()` y `pf()`, entregan percentiles y probabilidades acumuladas respectivamente.

# MÚLTIPLES VARIABLES ALEATORIAS

Los conceptos esenciales definidos para una variable aleatoria pueden ser extendidos a dos o más con la correspondiente distribución de probabilidades conjunta.

Considere por ejemplo dos variables aleatorias:

- $X$ , pluviosidad en cierta estación meteorológica
- $Y$ , nivel de un río.

Así, el evento  $(X = x, Y = y)$  se puede definir como  $(X = x \cap Y = y)$ , digamos evento conjunto, de igual forma el evento  $(X < x, Y < y)$  puede definirse como el evento conjunto  $(X < x \cap Y < y)$  definido para el espacio  $\Theta_{xy}$  de las variables aleatorias.



Para el par de variables aleatorias  $X$  e  $Y$  se define la función de distribución de probabilidad acumulada como

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$$

La cual satisface la axiomática fundamental de probabilidades:

- $F_{X,Y}(-\infty, -\infty) = 0.$
- $F_{X,Y}(-\infty, y) = 0.$
- $F_{X,Y}(x, -\infty) = 0.$
- $F_{X,Y}(x, +\infty) = F_X(x).$
- $F_{X,Y}(+\infty, y) = F_Y(y).$
- $F_{X,Y}(+\infty, +\infty) = 1.$

Si las variables aleatorias  $X$  e  $Y$  son discretas, la función de distribución de probabilidad conjunta es

$$p_{X,Y}(x, y) = P(X = x, Y = y)$$

siendo su función de distribución de probabilidad acumulada igual a

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y) = \sum_{x_i \leq x} \sum_{y_j \leq y} P(X = x_i, Y = y_j),$$

con  $(x_i, y_j) \in \Theta_{X,Y}$ .

# MÚLTIPLES VARIABLES ALEATORIAS

Ahora, si las variables aleatorias  $X$  e  $Y$  son continuas, la función de densidad de probabilidad conjunta se define como:

$$f_{X,Y}(x, y) dx dy = P(x < X \leq x + dx, y < Y \leq y + dy)$$

Entonces,

$$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(u, v) dv du.$$

Si las derivadas parciales existen, entonces

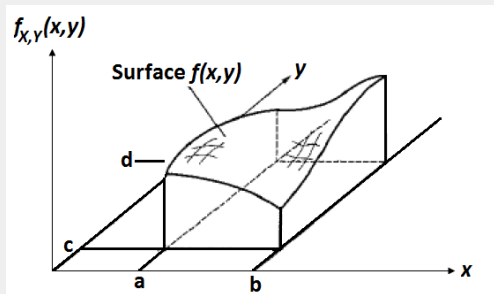
$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y)$$

# MÚLTIPLES VARIABLES ALEATORIAS

También, se puede observar que la siguiente probabilidad puede ser obtenida como

$$P(a < X \leq b, c < y \leq d) = \int_a^b \int_c^d f_{X,Y}(u, v) du dv$$

que representa el volumen bajo la superficie  $f_{X,Y}(x, y)$  como se muestra en la figura.



# MÚLTIPLES VARIABLES ALEATORIAS

Para variables aleatorias discretas  $X$  e  $Y$ , la probabilidad de ( $X = x$ ) puede depender de los valores que puede tomar  $Y$  (viceversa).

Con base a lo visto en probabilidades, se define la función de distribución de probabilidad condicional como:

$$p_{X|Y=y}(x) = P(X = x | Y = y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}, \quad p_Y(y) > 0$$

De manera similar, se tiene que

$$p_{Y|X=x}(y) = P(Y = y | X = x) = \frac{p_{X,Y}(x, y)}{p_X(x)}, \quad p_X(x) > 0$$

# MÚLTIPLES VARIABLES ALEATORIAS

La distribución marginal de una variable aleatoria se puede obtener aplicando el teorema de probabilidades totales.

Para determinar la distribución marginal de  $X$ ,  $p_X(x)$ , tenemos que

$$\begin{aligned} p_X(x) &= \sum_{y \in \Theta_Y} p_{X|Y=y}(x) \cdot p_Y(y) \\ &= \sum_{y \in \Theta_Y} p_{X,Y}(x, y) \end{aligned}$$

De la misma forma se tiene que

$$p_Y(y) = \sum_{x \in \Theta_X} p_{X,Y}(x, y)$$

# MÚLTIPLES VARIABLES ALEATORIAS

En el caso que ambas sean variables aleatorias continuas se define la función de densidad condicional de  $X$  dado que  $Y = y$  como

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x, y)}{f_Y(y)} \quad f_Y(y) > 0$$

De manera similar se tiene que

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x, y)}{f_X(x)} \quad f_X(x) > 0$$

Las respectivas marginales se obtienen como sigue:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$$
$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx$$

En el caso mixto, supongamos  $X$  discreta e  $Y$  continua, el calculo de las respectivas marginales es

$$p_X(x) = \int_{-\infty}^{\infty} p_{X|Y=y}(x) \cdot f_Y(y) dy$$
$$f_Y(y) = \sum_{x \in \Theta_X} f_{Y|X=x}(y) \cdot p_X(x)$$



Si ambas variables aleatorias son independientes, entonces se tiene que

$$p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y)$$

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$$

**Ejemplo 3.7:** Considere un autopista concesionada y un portico TAG en particular.

Datos históricos indica que el  $p \times 100\%$  de los automóviles adulteran su placa patente para no pagar TAG. Suponga que en promedio pasan en una hora  $\nu$  automóviles según un proceso de Poisson.

Proponga una distribución conjunta para las siguientes variables aleatorias:

- $X$ : Número de vehículos que pasan en una hora por el portico.
- $Y$ : Número de vehículos que pasan con patente adulterada en una hora por el portico.

Tenemos que

$$X \sim \text{Poisson}(\nu) \quad y \quad Y | X = x \sim \text{Binomial}(x, p)$$

Luego

$$\begin{aligned} p_{X,Y}(x, y) &= p_{Y|X=x}(y) \cdot p_X(x) \\ &= \binom{x}{y} p^y (1-p)^{x-y} \cdot \frac{\nu^x e^{-\nu}}{x!} \end{aligned}$$

$$\Theta_{X,Y} = \{(x, y) \mid x \in \mathbb{N}_0, y \in \mathbb{N}_0, y \leq x\}.$$

Por probabilidades totales se tiene que

$$Y \sim \text{Poisson}(\nu p)$$

## MÚLTIPLES VARIABLES ALEATORIAS

```
p.xy = function(x, y, p = 0.05, nu = 50){  
  n.r = length(x)  
  n.c = length(y)  
  M = matrix(NA, ncol = n.c, nrow = n.r)  
  for(i in 1:n.r){  
    M[i,] = dbinom(y, size = x[i], prob = p)*dpois(x[i],  
    lambda = nu)  
  }  
  M  
}
```

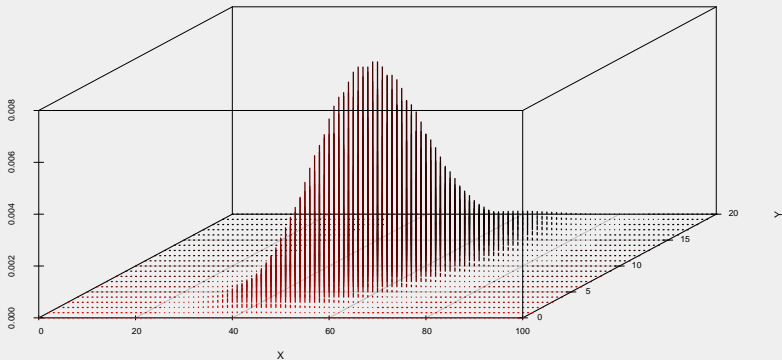
```
x = 0:100  
y = 0:20  
z = p.xy(x, y, p = 0.20, nu = 50)
```

# MÚLTIPLES VARIABLES ALEATORIAS

```
Z = c(z)
X = rep(x, length(y))
Y = sort(rep(y, length(x)))

install.packages("scatterplot3d")
library(scatterplot3d)
scatterplot3d(x = X, y = Y, z = Z, type = "h", lwd=2, pch
xlab = "X", ylab = "Y", zlab = "", highlight.3d=TRUE, ang
```

# MÚLTIPLES VARIABLES ALEATORIAS



**Ejemplo 3.8:** Sean  $X$  e  $Y$  variables aleatorias continuas, relacionadas de la siguiente manera:

$$Y \sim \text{Gamma}(k, \nu) \quad y \quad X | Y = y \sim \text{Uniforme}(0, y)$$

Luego

$$f_{X,Y}(x, y) = \frac{\nu^k}{\Gamma(k)} y^{k-1} e^{-\nu y} \cdot \frac{1}{y}, \quad 0 \leq x \leq y$$

# MÚLTIPLES VARIABLES ALEATORIAS

```
f.xy = function(x, y, k = 3, nu = 2){  
  n.r = length(x)  
  n.c = length(y)  
  M = matrix(NA, ncol = n.c, nrow = n.r)  
  for(i in 1:n.r){  
    M[i,] = dgamma(y, rate = nu, shape = k)*dunif(x[i], 0, y)  
  }  
  M  
}  
x = seq(0,5,.05)  
y = seq(0,5,.05)  
z = f.xy(x, y)  
install.packages("rgl")  
library(rgl)  
rgl.surface(x = x, y = z, z = y, color = "red", back="lin
```



**Ejemplo 3.9:** Dos variables aleatorias  $X$  e  $Y$  tienen distribución conjunta Normal-Bivariada si su función de densidad conjunta está dada por:

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \times \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right]\right\}$$

A partir de esto se puede deducir que

$$X \sim \text{Normal}(\mu_X, \sigma_X), \quad Y \sim \text{Normal}(\mu_Y, \sigma_Y)$$

$$Y | X = x \sim \text{Normal}\left(\mu_Y + \frac{\rho\sigma_Y}{\sigma_X}(x - \mu_X), \sigma_Y\sqrt{(1-\rho^2)}\right)$$

# MÚLTIPLES VARIABLES ALEATORIAS

```
f.xy = function(x, y, mu.x = 0, mu.y = 0, s.x = 1, s.y =  
rho = 0){  
  n.r = length(x)  
  n.c = length(y)  
  M = matrix(NA, ncol = n.c, nrow = n.r)  
  for(i in 1:n.r){  
    M[i,] = dnorm(x[i], mean = mu.x, sd = s.x) * dnorm(y, mea  
+ rho*s.y*(x[i]-mu.x)/s.x, sd = s.y*sqrt(1-rho^2))  
  }  
  M  
}  
x = seq(-5,5,0.1)  
y = seq(-5,5,0.1)  
z = f.xy(x, y, rho = 0)  
rgl.surface(x = x, y = z*10, z = y, color = "red", back="
```

# MÚLTIPLES VARIABLES ALEATORIAS

Cuando hay dos variable aleatoria  $X$  e  $Y$ , puede haber una relación entre las variables.

En particular, la presencia o ausencia de relación estadística lineal se determina observando el primer momento conjunto de  $X$  e  $Y$  definido como

$$E(X Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \cdot f_{X,Y}(x, y) dx dy$$

Si  $X$  e  $Y$  son estadísticamente independientes, entonces

$$E(X Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \cdot f_X(x) \cdot f_Y(y) dx dy = E(X) \cdot E(Y)$$

La covarianza corresponde al segundo momento central y se define como:

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = E(X \cdot Y) - \mu_X \cdot \mu_Y$$

Si  $X$  e  $Y$  son estadísticamente independientes, entonces

$$\text{Cov}(X, Y) = 0$$

El significado físico de la covarianza se puede inferir de la ecuación.

Si la  $\text{Cov}(X, Y)$  es grande y positivo, los valores de  $X$  e  $Y$  tienden a ser grandes o pequeños en relación a sus respectivos medias, mientras que si la  $\text{Cov}(X, Y)$  es grande y negativo, los valores de  $X$  tienden a ser grandes con respecto a su media, mientras que los de  $Y$  tienden a ser pequeños.

El significado físico de la covarianza se puede inferir de la ecuación anterior:

- Si  $\text{Cov}(X, Y)$  es grande y positiva, los valores de  $X$  e  $Y$  tienden a ser grandes (o pequeños) en relación a sus respectivos medias.
- Si  $\text{Cov}(X, Y)$  es grande y negativo, los valores de  $X$  tienden a ser grandes con respecto a su media, mientras que los de  $Y$  tienden a ser pequeños y viceversa.
- Si  $\text{Cov}(X, Y)$  es pequeña o cero, la relación (lineal) entre los valores de  $X$  e  $Y$  es poca o nula, o bien la relación es No lineal.

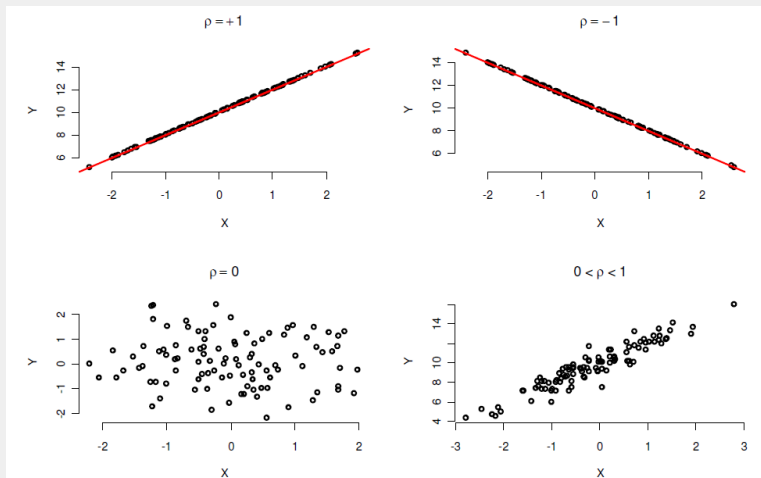
La covarianza mide el grado de asociación lineal entre dos variables, pero es preferible su normalización llamada correlación para poder cuantificar la magnitud de la relación.

La Correlación esta definida como:

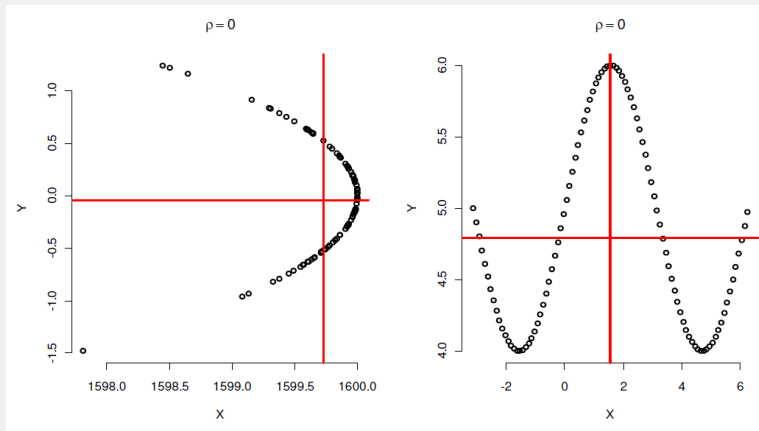
$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y}$$

Este coeficiente toma valores en el intervalo  $(-1, 1)$ .

# MÚLTIPLES VARIABLES ALEATORIAS



# MÚLTIPLES VARIABLES ALEATORIAS





# ESPERANZA      CONDICIONAL      Y PREDICCIÓN

## ESPERANZA CONDICIONAL Y PREDICCIÓN (RICE SEC 4.4)

El valor esperado de una variable aleatoria  $Y$  condicionado a la realización  $x$  de una variable aleatoria  $X$  esta dado por

$$E(Y | X = x) = \begin{cases} \sum_{y \in \Theta_{Y|X=x}} y \cdot P(Y = y | X = x), & \text{caso discreto} \\ \int_{y \in \Theta_{Y|X=x}} y \cdot f_{Y|X=x}(y) dy, & \text{caso continuo} \end{cases}$$

## ESPERANZA CONDICIONAL Y PREDICCIÓN (RICE SEC 4.4)

Por otra parte, para una función de  $Y$ , llamemos  $h(Y)$ , el valor esperado condicional esta dado por

$$E[h(Y) | X = x] = \begin{cases} \sum_{y \in \Theta_{Y|X=x}} h(y) \cdot P(Y = y | X = x), & \text{caso discreto} \\ \int_{y \in \Theta_{Y|X=x}} h(y) \cdot f_{Y|X=x}(y) dy, & \text{caso continuo} \end{cases}$$

# ESPERANZA CONDICIONAL Y PREDICCIÓN (RICE SEC 4.4)

## Ejemplo 3.10

Considere un proceso de Poisson cuyo número de eventos esperado en  $[0, 1]$  es  $\lambda$ . Si  $N$  y  $X$  son variables aleatorias que determinan la cantidad de eventos en los intervalos  $[0, 1]$  y  $[0, p]$  respectivamente, con  $0 < p < 1$ , muestre que  $E(X | N = n) = np$ .

# ESPERANZA CONDICIONAL Y PREDICCIÓN (RICE SEC 4.4)

## Ejemplo 3.11

Sean  $X$  e  $Y$  dos variables aleatorias cuya distribución de probabilidad conjunta está determinada por una Normal bivariada de parámetros  $\mu_X, \mu_Y, \sigma_X, \sigma_Y$  y  $\rho$ .

Muestre que

$$E(Y | X = x) = \mu_Y + \rho(x - \mu_X)\sigma_Y/\sigma_X \quad \text{y} \quad \text{Var}(Y | X = x) = \sigma_Y^2(1 - \rho^2)$$

## Teorema de probabilidades totales para valor esperado

$$E(X) = \begin{cases} \sum_{y \in \Theta_Y} \left[ \sum_{x \in \Theta_{X|Y=y}} x \cdot p_{X|Y=y}(x) \right] p_Y(y), & \text{Caso Discreto-Discreto} \\ \int_{y \in \Theta_Y} \left[ \int_{x \in \Theta_{X|Y=y}} x \cdot f_{X|Y=y}(x) dx \right] f_Y(y) dy, & \text{Caso Continuo-Continuo} \\ \int_{y \in \Theta_Y} \left[ \sum_{x \in \Theta_{X|Y=y}} x \cdot p_{X|Y=y}(x) \right] f_Y(y) dy, & \text{Caso Discreto-Continuo} \\ \sum_{y \in \Theta_Y} \left[ \int_{x \in \Theta_{X|Y=y}} x \cdot f_{X|Y=y}(x) dx \right] p_Y(y), & \text{Caso Continuo-Discreto} \end{cases}$$

# ESPERANZA CONDICIONAL Y PREDICCIÓN (RICE SEC 4.4)

## Teoremas Esperanza Iterada

$$E(Y) = E[E(Y | X)]$$

$$\text{Var}(Y) = \text{Var}[E(Y | X)] + E[\text{Var}(Y | X)]$$

## ESPERANZA CONDICIONAL Y PREDICCIÓN (RICE SEC 4.4)

Predecir el valor de una variable aleatoria a partir de otra es el propósito de esta sección.

Consideremos primero la siguiente situación: “Predecir la realización de una variable aleatoria  $Y$ ”. El “mejor” valor  $c$  para predecir la realización de  $Y$  se puede obtener minimizando el error cuadrático medio definido como

$$\text{ECM} = E[(Y - c)^2]$$

Mostrar que la constante  $c$  que minimiza el ECM es  $E(Y)$ .



## ESPERANZA CONDICIONAL Y PREDICCIÓN (RICE SEC 4.4)

Si ahora queremos predecir  $Y$  basado en una función de una variable aleatoria  $X$ , llamemos  $h(X)$ , que minimice el error cuadrático medio definido como

$$\begin{aligned}\text{ECM} &= E \left\{ [Y - h(X)]^2 \right\} \\ &= E \left( E \left\{ [Y - h(X)]^2 \mid X \right\} \right)\end{aligned}$$

Entonces, la función  $h(X)$  que minimiza ECM necesariamente debe corresponder a  $E(Y \mid X)$ .

Por ejemplo, si  $X$  e  $Y$  distribuyen conjuntamente según una Normal bivariada, entonces el mejor predictor  $Y$  basado en  $X$  es una función lineal dada por

$$E(Y \mid X) = \left( \mu_Y - \mu_X \frac{\rho \sigma_Y}{\sigma_X} \right) + X \frac{\rho \sigma_Y}{\sigma_X}$$