

## Glosario de funciones más utilizadas en Fathom

Se listan aquí algunas de las funciones más usadas en Fathom, su estructura, sus resultados y su ubicación. Existen más funciones, pero se omiten aquí dado que no son tan socorridas a un nivel básico del manejo del paquete. También se anotan algunos términos especiales de Fathom.

**Attribute.**- Con este nombre Fathom se refiere a los nombres empleados para las variables.

En el menú Functions de la ventana de fórmula se encuentran, entre otras, las siguientes funciones, según su clasificación:

### 1. Arithmetic:

**abs(*a*).**- Genera el valor absoluto de una cantidad *a*.

**combinations (*n,r*).**- Genera el número de combinaciones posibles de *r* elementos seleccionados de entre *n* totales.

**exp(*a*).**- Genera el resultado de elevar el número de Napier a la potencia *a*.

**ln(*a*).**- Genera el logaritmo natural de la cantidad *a*.

**log(*a*).**- Genera el logaritmo en base diez de la cantidad *a*.

**modulo(*a,b*).**- Genera el residuo obtenido del cociente *a/b*.

**round(*a*).**- Redondea la cantidad racional *a* al entero más cercano.

**sqrt(*a*).**- Genera la raíz cuadrada de la cantidad *a*.

**trunc(*a*).**- Genera el entero resultante de truncar la parte decimal del número racional *a*.

### 2. Condicional:

**if(*condición*)**  $\left\{ \begin{array}{l} ? \\ ? \end{array} \right.$ .- Permite elegir entre dos opciones dependiendo de si se cumple o

no la *condición* establecida entre los paréntesis. Las opciones se anotan delante del corchete y en lugar de los signos de interrogación, colocando siempre en primera instancia la opción a tomar si la *condición* establecida efectivamente de cumple, y en segunda instancia la opción a tomar si dicha *condición* no se cumple. La condición puede ser una igualdad, una desigualdad o incluso alguna función de Fathom.

**switch(*atributo*)**  $\left\{ \begin{array}{l} () : ? \\ else : ? \end{array} \right.$ .- Permite elegir entre muchas posibilidades, dependiendo

de diferentes condiciones para un *atributo*; estas condiciones se anotan en los paréntesis situados junto a los signos de interrogación, los cuales deberán de sustituirse por la acción a realizar su efectivamente se observa el valor dado para el *atributo*. Por defecto, se muestran solo dos alternativas en la función, pero tras escribir la acción para primera condición, pueden agregarse más alternativas presionando la tecla **Insert**. La opción *else* solo aparecerá para la última opción, la cual se realizará solo si ninguna de las condiciones antes establecidas se cumpliera.

### 3. Distributions:

Este tipo de funciones tienen todas la misma estructura, y generan valores asociados a diferentes distribuciones de probabilidad. Las distribuciones con las que Fathom trabaja son Beta, Binomial, Cauchy, Ji-Cuadrada, Exponencial, F, Gama, Geométrica, Hipergeométrica, Normal, Poisson, t de Student, Uniforme y Uniforme de lattice. Como ejemplo del manejo de las funciones asociadas a estas distribuciones tomaremos el caso de la distribución Normal, aunque se recomienda consultar en la parte baja de la misma ventana de funciones qué parámetros deben incluirse en cada caso en particular:

`normalCumulative(x,m,s)`.- Genera la probabilidad de obtener un valor menos o igual a  $x$ , bajo una distribución de probabilidad Normal, con media  $m$  y desviación estándar  $s$ . Los valores de  $m$  y  $s$  por defecto son respectivamente de uno y cero.

`normalDensity(x,m,s)`.- Genera la densidad de probabilidad de obtener un valor  $x$ , bajo una distribución de probabilidad Normal, con media  $m$  y desviación estándar  $s$ .

`normalQuantile(c,m,s)`.- Genera el valor de la variable normalmente distribuida, con media  $m$  y desviación estándar  $s$ , para la cual la probabilidad acumulada es  $c$ .

### 4. Random:

En este tipo de funciones pueden generarse números aleatorios, algunos de ellos

acordes a algunas de las distribuciones de probabilidad que se mencionaron en el apartado anterior. La sintaxis es muy semejante en estos casos, pero de nuevo es necesario revisar qué parámetros deben indicarse en la estructura de la función. El resto de las funciones Random son:

`random(a,b)`.- Genera un número aleatorio entre  $a$  y  $b$ . Si no se establecen estos parámetros, y la función se escribe simplemente como `random`, se generará un número aleatorio entre cero y uno.

`randomInteger(a,b)`.- Genera un número aleatorio entero entre  $a$  y  $b$ .

`randomPick(a,b)`.- Genera una selección aleatoria de uno entre los valores  $a$  y  $b$ , siendo estos caracteres alfanuméricos. Pueden incluirse hasta diez caracteres como argumentos de esta función.

### 5. Statistical:

a) One Attribute.- En este apartado se encuentran algunas funciones estadísticas manejadas en el caso de trabajar con datos univariados:

`count(atributo)`.- Cuenta el número de observaciones para un *atributo*. Igualmente puede darse una condición al *atributo*, como `count(atributo<a)`, con lo que se contará el número de observaciones que cumplen la condición para el *atributo*. Si se omite el argumento y se escribe `count`, entonces se generará el total de observaciones en la colección.

`first(atributo)`.- Indica el primer valor que se observó para el *atributo* indicado.

`iqr(atributo)`.- Genera el rango intercuartílico del *atributo*.

`last(atributo)`.- Indica el último valor observado para el *atributo* indicado.

`max(atributo)`.- Indica el valor máximo observado del *atributo*.  
`mean(atributo)`.- genera la media aritmética del *atributo*. Es posible insertar alguna condición, con la estructura `mean(atributo1, atributo2=a)`, en donde se generará la media aritmética de las observaciones del *atributo1* que mostraron el valor *a* en el *atributo2*.

`median(atributo)`.- Genera la mediana del *atributo*.

`min(atributo)`.- Indica el valor mínimo observado del *atributo*.

`percentile(a,atributo)`.- Genera el valor de *a*-ésimo percentil para el *atributo*.

`popStdDev(atributo)`.- Genera la desviación estándar poblacional del *atributo*.

`popVariance(atributo)`.- Genera la variancia poblacional del *atributo*.

`proportion(atributo=a)`.- Genera la proporción de casos en los que el *atributo* toma el valor *a*, que puede ser numérico o alfanumérico.

`Q1(atributo)`.- Genera el valor del primer cuartil para el *atributo*.

`Q3(atributo)`.- Genera el valor del tercer cuartil del *atributo*.

`s(atributo)`.- Genera la desviación estándar muestral del *atributo*. Igualmente puede obtenerse con las funciones `sampleStdDev(atributo)` y `stdDev(atributo)`.

`sampleVariance(atributo)`.- Genera la variancia muestral del *atributo*. Igualmente puede obtenerse con la función `variance(atributo)`.

`sum(atributo)`.- Genera la suma de los valores numéricos del *atributo*.

`uniqueValues(atributo)`.- Indica el número de valores diferentes que toma el *atributo*.

#### b) Transformations:

`next(atributo)`.- Genera el siguiente valor para el *atributo*. Puede escribirse `next(atributo,a)` en donde *a* es el valor a generar si no hay un siguiente valor del *atributo*.

`popZScore(atributo)`.- Indica el número de desviaciones estándar de distancia a la media para cada valor del *atributo*, asumiendo que se trabaja con una población.

`prev(atributo)`.- Genera el valor previo para el *atributo*. Puede escribirse `prev(atributo,a)` en donde *a* es el valor a generar si no hay un valor previo del *atributo*.

`sampleZScore(atributo)`.- Indica el número de desviaciones estándar de distancia a la media para cada valor del *atributo*, asumiendo que se trabaja con una muestra.

#### c) Two Attributes:

`correlation(a,b)`.- Genera el coeficiente de correlación entre los atributos *a* y *b*.

`covariance(a,b)`.- Genera la covariancia entre los atributos *a* y *b*.

`rSquared(a,b)`.- Genera el coeficiente de determinación entre los atributos *a* y *b*.

#### 6. Special:

`caseIndex`.- Indica el número de casos en una colección.

En la ventana Inspector que se utiliza al trabajar con muestras generadas por el usuario se tiene la función `Until`, con la que puede establecerse una condición, por ejemplo utilizando alguna función, para detener el proceso de muestreo.

Tomado de Paquete didáctico de Estadística y Probabilidad I, de Antuna y Hernández.