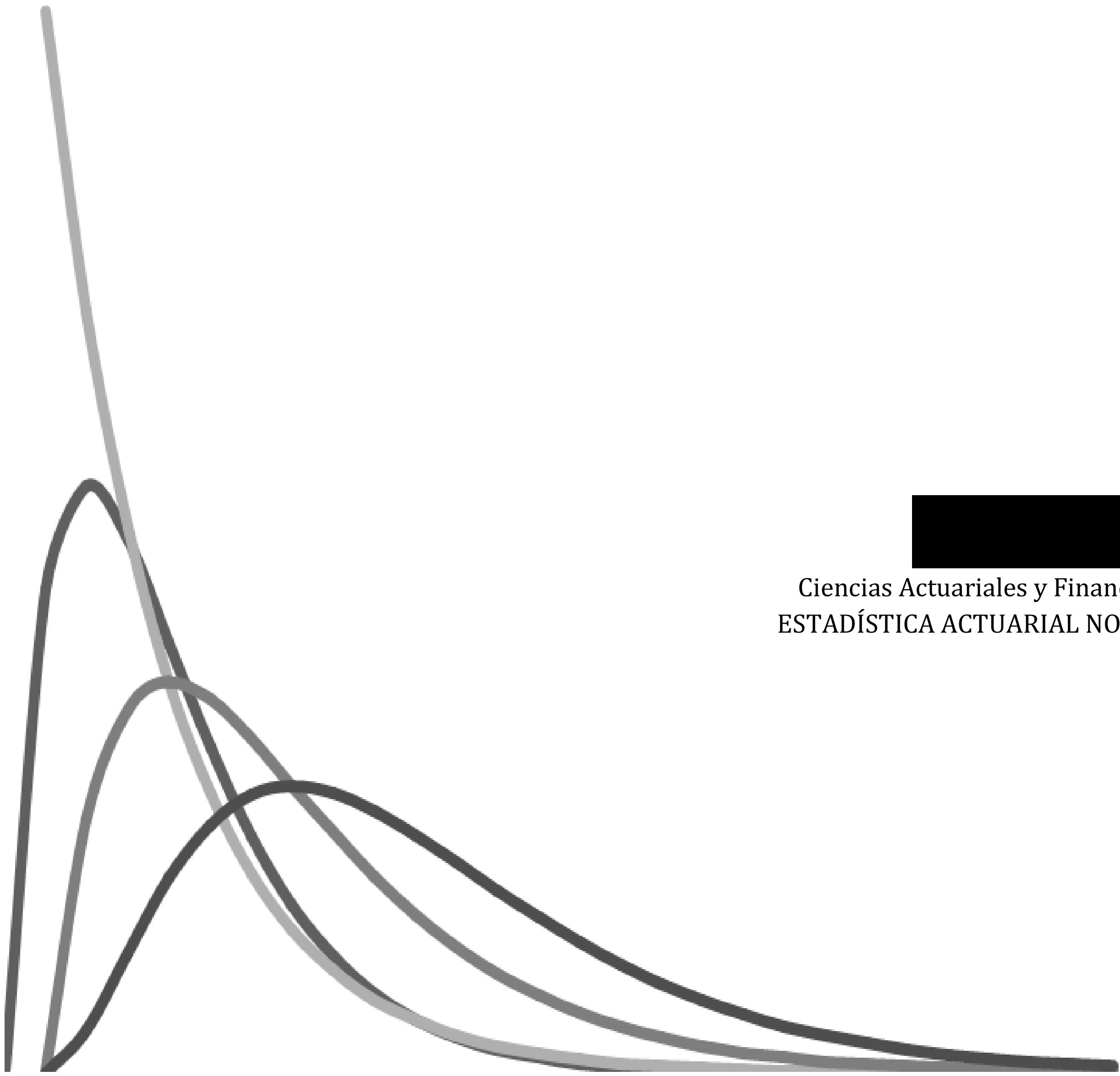
DISTRIBUCIÓN GAMMA



[redacted]
Ciencias Actuariales y Financieras
ESTADÍSTICA ACTUARIAL NO VIDA

La distribución Gamma

Es una distribución adecuada para modelizar el comportamiento de variables aleatorias continuas con asimetría positiva. Es decir, variables que presentan una mayor densidad de sucesos a la izquierda de la media que a la derecha. En su expresión se encuentran dos parámetros, siempre positivos, (α) y (β) de los que depende su forma y alcance por la derecha, y también la función Gamma Γ(α), responsable de la convergencia de la distribución.

Los parámetros de la distribución

El primer parámetro (α) situa la máxima intensidad de probabilidad y por este motivo en algunas fuentes se denomina “la forma” de la distribución: cuando se toman valores próximos a cero aparece entonces un dibujo muy similar al de la distribución exponencial. Cuando se toman valores más grandes de (α) el centro de la distribución se desplaza a la derecha y va apareceiendo la forma de una campana de Gauss con asimetría positiva. Es el segundo parámetro (β) el que determina la forma o alcance de esta asimetría positiva desplazando la densidad de probabilidad en la cola de la derecha. Para valores elevados de (β) la distribución acumula más densidad de probabilidad en el extremo derecho de la cola, alargando mucho su dibujo y dispersando la probabilidad a lo largo del plano. Al dispersar la probabilidad la altura máxima de densidad de probabilidad se va reduciendo; de aquí que se le denomine “escala”. Valores más pequeños de (β) conducen a una figura más simétrica y concentrada, con un pico de densidad de probabilidad más elevado.

Una forma de interpretar (β) es “tiempo promedio entre ocurrencia de un suceso”. Relacionándose con el parámetro de la Poisson como β=1/λ. Alternativamente λ será el ratio de ocurrencia: λ=1/β.
La expresión β^{-1} también será necesaria más adelante para poder llevar a cabo el desarrollo matemático.

Relación con otras distribuciones

Si se tiene un parámetro α de valores elevados y β pequeña, entonces la función Gamma converge con la distribución normal. De media $\mu = \alpha * \beta$, y varianza $\sigma^2 = \alpha * \beta^2$.
Cuando $\alpha = 1$ y $\beta = 0$ la distribución Gamma es exactamente la distribución exponencial con parámetro (α=1).
Cuando la proporción entre parámetros es $[\alpha = \frac{v}{2} : \beta = v]$ entonces la variable aleatoria se distribuye como una Chi-cuadrado con v grados de libertad.
Si α=1, entonces se tiene la distribución exponencial negativa de parámetro λ=1/β.

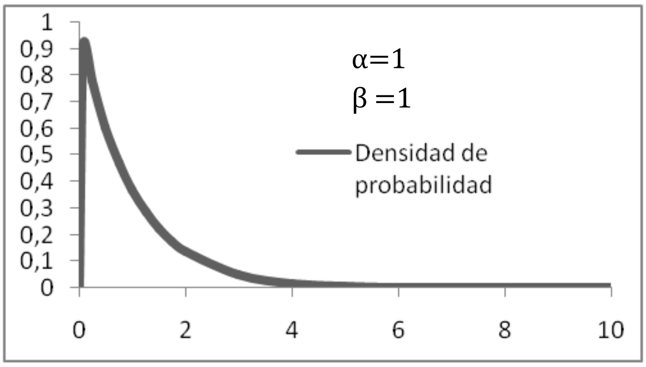
Ventajas

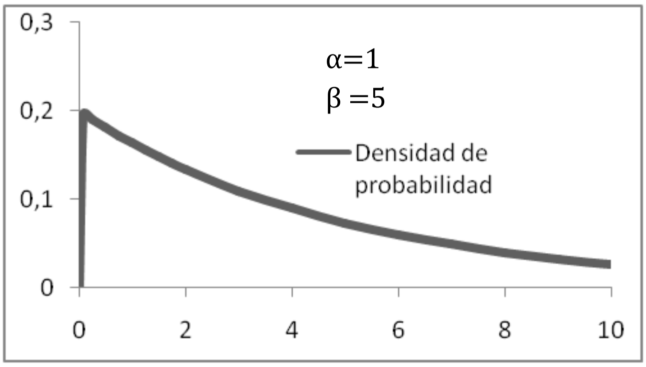
De esta forma, la distribución Gamma es una distribución flexible para modelizar las formas de la asimetría positiva, de las más concentradas y puntiagudas, a las más dispersas y achatadas. Como ejemplos de variables que se comportan así:
- Número de individuos involucrados en accidentes de tráfico en el área urbana: es más habitual que la mayoría de partes abiertos den la proporción de 1 herido por vehículo, que otras proporciones superiores.
- Altura a la que se inician las precipitaciones; sucede de forma más habitual precipitaciones iniciadas a una altura baja, que iniciadas a gran altitud.
- Tiempo o espacio necesarios para observar X sucesos que siguen una distribución de Poisson.
- Distribución de la finura de fibras de lana: la mayoría presentan una menor finura que unas pocas fibras más gruesas.

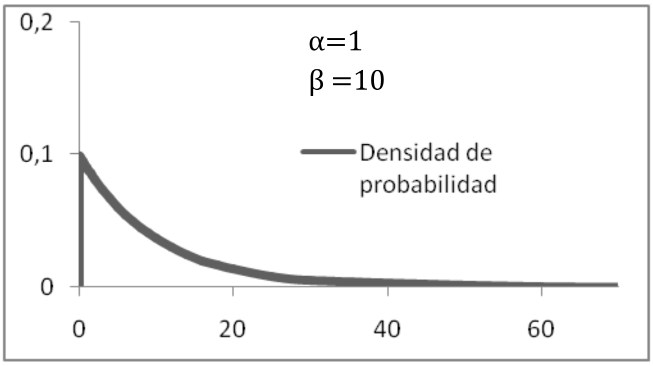
Inconvenientes

Problemas en la complejidad de algunos cálculos, especialmente respecto a la función Gamma cuando el parámetro α es un valor no entero. También problemas de cálculo en la estimación de los parámetros muestrales. Ambos inconvenientes se pueden abordar satisfactoriamente con ordenador.

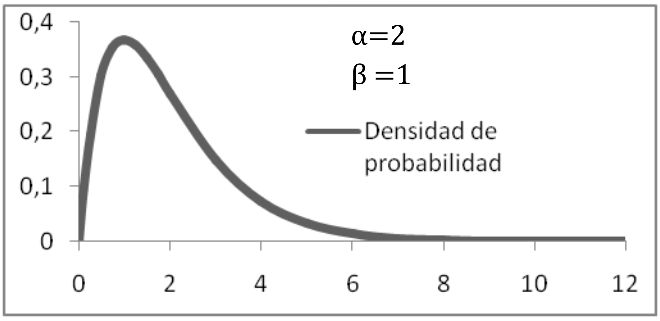
Para valorar la evolución de la distribución al variar los parámetros se tienen los siguientes gráficos. Primero se comprueba que para α=1 la distribución tiene similitudes con la exponencial.

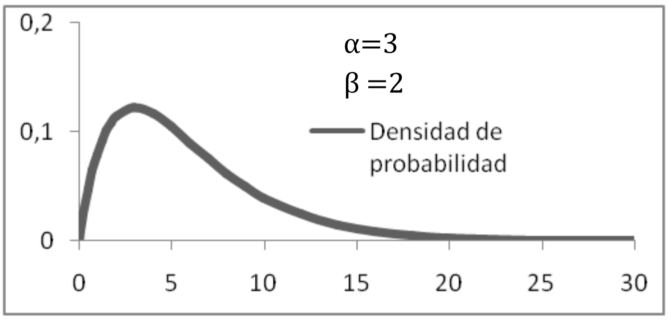


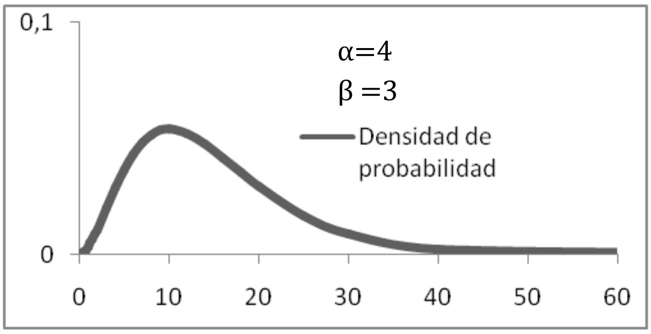




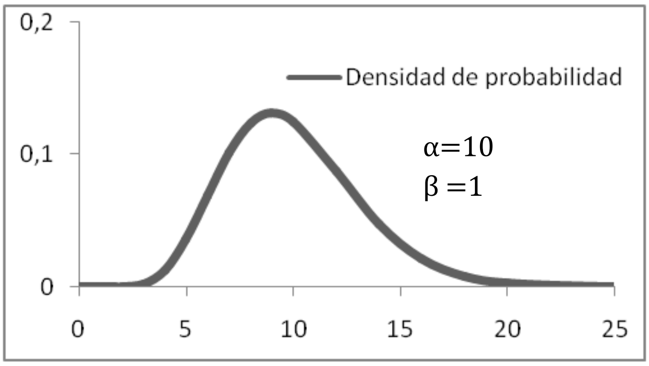
Si ahora se hace variar el parámetro alfa,







Y para valores altos de α y pequeños de β, se observa la convergencia con la normal,



Como ya se ha indicado, la expresión de la distribución Gamma incluye la propia función Gamma $\Gamma(\alpha)$, que para valores enteros de alpha se ha demostrado que $\Gamma(\alpha) = (\alpha - 1)!$ En este caso la distribución Gamma se conoce como "distribución de Erlang".

Valor α	Función Gamma
1	1
2	1
3	2
4	6
5	24
6	120
7	720
8	5040
9	40320
10	362880
11	3628800
12	39916800

Para valores no enteros, **el valor de la función Gamma** se obtiene a partir de,

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} * e^{-x} dx$$

La **función de densidad** de la distribución Gamma es,

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} * x^{\alpha-1} * e^{-\frac{x}{\beta}}$$

donde x>0 y β,α son parámetros positivos.

Se puede ver que para α=1 la función de densidad será,

$$f(x) = \frac{1}{1 * 1} * x^0 * e^{-\frac{x}{1}} = e^{-x}$$

lo que significa que una distribución Gamma de parámetro α=1 y β =0 es una distribución exponencial de parámetro α=1.

Se demuestra que f(x) es una función de densidad porque para f(x)≥0 y haciendo el cambio de variable $\mathrm{h} = \frac{1}{\beta}$

$$\int_0^{\infty} f(x)dx = \frac{\mathrm{h}^{\alpha}}{\Gamma(\alpha)} * \int_0^{\infty} x^{\alpha-1} * e^{-x\mathrm{h}} * dx = \frac{h^{\alpha}}{\Gamma(\alpha)} * \frac{\Gamma(\alpha)}{h^{\alpha}} = 1$$

La **función de distribución** es,

$$F(x) = P(X \leq x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} * \int_0^{x} u^{\alpha-1} * e^{-\frac{u}{\beta}} * du$$

La **función característica** es, teniendo en cuenta de nuevo que $\mathrm{h} = \frac{1}{\beta}$

$$\ell_x(t) = \int_{-\infty}^{+\infty} e^{itx} f(x)dx = \int_0^{+\infty} e^{itx} \frac{h^{\alpha}}{\Gamma(\alpha)} * x^{\alpha-1} * e^{-hx} * dx = \frac{h^{\alpha}}{\Gamma(\alpha)} * \int_0^{\infty} x^{\alpha-1} * e^{-(\mathrm{h}-it)x} dx =$$

$$= \left\{ \frac{\Gamma(\alpha)}{(\mathrm{h} - i * t)^{\alpha}} = \int_0^{+\infty} x^{\alpha-1} * e^{-(\mathrm{h}-it)x} dx \right\} = \ell_x(t) = \frac{\mathrm{h}^{\alpha}}{\Gamma(\alpha)} * \frac{\Gamma(\alpha)}{(\mathrm{h} - i * t)^{\alpha}} = \left(\frac{\mathrm{h}}{\mathrm{h} - it}\right)^{\alpha} = \left(\frac{1}{1 - \frac{it}{\mathrm{h}}}\right)^{\alpha} =$$

$$\ell_x(t) = \left(1 - \frac{it}{\mathrm{h}}\right)^{-\alpha}$$

y volviendo a β=1/h,

$$\ell_x(t) = (1 - \beta * i * t)^{-\alpha}$$

La **esperanza matemática** será, siendo $\mathrm{h} = \frac{1}{\beta}$

$$E(X) = \int_0^{+\infty} x * f(x) * dx = \frac{h^{\alpha}}{\Gamma(\alpha)} \int_0^{\infty} x * x^{\alpha-1} * e^{-hx} * dx = \frac{h^{\alpha}}{\Gamma(\alpha)} \int_0^{\infty} x^{\alpha} * e^{-hx} * dx =$$

$$= \frac{h^{\alpha}}{\Gamma(\alpha)} * \frac{\Gamma(\alpha+1)}{h^{\alpha+1}} = \frac{1}{\mathrm{h}} * \frac{\alpha * \Gamma(\alpha)}{\Gamma(\alpha)} = \frac{\alpha}{\mathrm{h}}$$

Volviendo a β;

$$E(X) = \alpha\beta$$

La **varianza** será,

$$V(X) = E(X^2) - \left(E(X)\right)^2$$

donde otra vez $h = \frac{1}{\beta}$

$$E(X^2) = \int_0^{\infty} x^2 f(x)dx = \frac{h^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^2 * x^{\alpha-1} * e^{-hx} * dx = \frac{h^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^{\alpha+1} * e^{-hx} * dx =$$

$$\frac{h^\alpha}{\Gamma(\alpha)} * \frac{\Gamma(\alpha+2)}{h^{\alpha+2}} = \frac{1}{h^2} * \frac{(\alpha+1) * \Gamma(\alpha)}{\Gamma(\alpha)} = \frac{\alpha(\alpha+1)}{h^2}$$

$$V(X) = E(X^2) - \left(E(X)\right)^2 = \frac{\alpha(\alpha+1)}{h^2} * \left(\frac{\alpha}{h}\right)^2 = \frac{\alpha^2 + \alpha - \alpha^2}{h^2} = \frac{\alpha}{h^2}$$

Y volviendo a β;

$$V(X) = \alpha\beta^2 \quad \rightarrow desv.típica = D(X) = \beta\sqrt{\alpha}$$

Si se buscan estimadores de los parámetros de la distribución, el Método de Máxima Verosimilitud es el más adecuado ya que goza de más propiedades que el estimador obtenido por el método de los momentos (Chico, 2010). Sin embargo la máxima verosimilitud conduce a un sistema de ecuaciones que no se puede resolver de forma analítica, y se necesita recurrir al método numérico, por ejemplo de Newton-Raphson.

$$-ln\beta - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^{n} lnx_i = 0$$

$$\alpha = \frac{\sum_{i=1}^{n} x_i}{n * \beta}$$

Resolución de probabilidades de la distribución Gamma

En R se pueden usar las órdenes >dgamma y >pgamma que devuelven resultados de la función de densidad y de densidad acumulada respectivamente.

>dgamma(x, "parámetro (α)", scale = "parámetro (β); escala")
Y devuelve el valor de la densidad de probabilidad en cuando la variable aleatoria tiene valor = x

>pgamma(x, "parámetro (α)", scale = "parámetro (β); escala"; lower.tail = T)
Devuelve el valor de probabilidad acumulada hasta el valor x. Con lower.tail = F devuelve la probabilidad por la derecha.

Si en lugar de parámetro escala se está usando el parámetro ratio se escribirá;
>dgamma(x, "parámetro (α)", ratio = "parámetro (β); escala")
>pgamma(x, "parámetro (α)", ratio = "parámetro (β); escala"; lower.tail = T)

En excel la función es, para la intensidad de probabilidad
=DISTR.GAMMA(x;"parámetro (α);"parámetro escala (β)";FALSO)
Y para la probabilidad acumulada,
=DISTR.GAMMA(x;"parámetro (α);"parámetro escala (β)";VERDADERO)
Pero no permite diferenciar entre "escala" o "ratio".

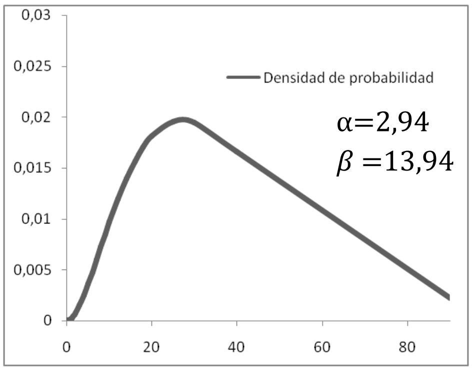
Ejemplos de distribución Gamma

Ejemplo 1.

En un estudio de la guardia urbana de Barcelona se toma una distribución gamma para modelizar el número de víctimas en accidentes de tráfico. Como es más habitual la proporción de 1 ocupante por vehículo siniestrado, y es más rara la probabilidad de 4 ó 5 ocupantes por vehículo siniestrado, se crea una distribución gamma para modelizar el número de víctimas por accidente de tráfico. El 38% de la distribución lo acumula la proporción 1 accidentado por accidente, el 36% 2:1, 16% la 3:1, 6% el 4:1 y finalmente un 3% para 5:1. La media del modelo es 1,5 víctimas por accidente, pero no indican el valor de los parámetros α y β tomados en cuenta.
(*Análisis del envío de equipos de emergencia para accidentes en las Rondas de Barcelona, pág. 40*)

Ejemplo 2.

También en el ámbito de la siniestralidad viaria, en un estudio de la ciudad de Medellín, Colombia, se usa la distribución Gamma para obtener la distribución de probabilidad de la variable aleatoria "edad de fallecimiento en accidentes de tráfico". En este caso explican que se asignaron los parámetros α y β "a ojo". El mejor resultado es el que parece minimizar los errores cuadráticos medios después de varias asignaciones. Finalmente obtienen α=2,94 y β =13,94.
(*Panorama de la accidentalidad vial en Medellín en 1999, pág 14*)



Ejemplo 3. Relación entre la distribución gamma y procesos de Poisson

A una centralita de teléfonos llegan 12 llamadas por minuto, siguiendo una distribución de Poisson. ¿Cuál es la probabilidad de que en menos de 1 minuto lleguen 8 llamadas?

Si $\lambda = 12 \frac{llamadas}{minuto}$ entonces,

$$\beta = \frac{1\ minuto}{12\ llamadas}$$

$$P(X < 1) = \frac{1}{\frac{1}{12}^{1} * 2} * \int_0^1 u^{2-1} * e^{-\frac{u}{1/12}} du$$

Resolviendo en R,

```
> pgamma(1, 8, scale = 1/12, lower.tail = T)
[1] 0.9104955
>
```

Existe un 91,05% de probabilidades de recibir 8 llamadas en un plazo de tiempo de menos de 1 minuto.

Ejemplo 4.
Si un componente eléctrico falla una vez cada 5 horas, ¿cuál es el tiempo medio que transcurre hasta que fallan dos componentes? ¿Cuál es la probabilidad de que transcurran 12 horas antes de que fallen los dos componentes?

Si $\lambda = \frac{1}{5}\frac{fallo}{horas}$ entonces, $\beta = \frac{1}{\lambda} = \frac{1}{\frac{1}{5}} =$ 5 horas/fallo

$$E(X) = \alpha * \beta = 2 * 5 = 10\ horas$$

$$P(X > 12) = \frac{1}{5^2 * 2} * \int_0^{12} u^{2-1} * e^{-\frac{u}{1/12}} du$$

Resolviendo en R
> pgamma(12, 2, scale = 5, lower.tail = F)
[1] 0.3084410
>

En un 30,84% de las situaciones pasarán 12 horas hasta que fallen dos componentes.

Ejemplo 5.
En una ciudad se observa que el consumo diario de energía (en millones de kilowatt-hora) es una variable aleatoria que sigue una distribución gamma con parámetros α= 3 y β =2. Si la planta de energía que suministra a la ciudad tiene una capacidad diaria de generar un máximo de 12, ¿cuál es la probabilidad de que haya un día donde no se pueda satisfacer la demanda?

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} * x^{\alpha-1} * e^{-\frac{x}{\beta}}$$

donde,

$$\alpha = 3 \quad \beta = 2$$
$$\Gamma(\alpha) = (n-1)! = \Gamma(3) = (3-1)! = 2$$

de forma que,

$$f(x) = \frac{1}{2^3 * 2} * x^2 * e^{-\frac{x}{2}}$$

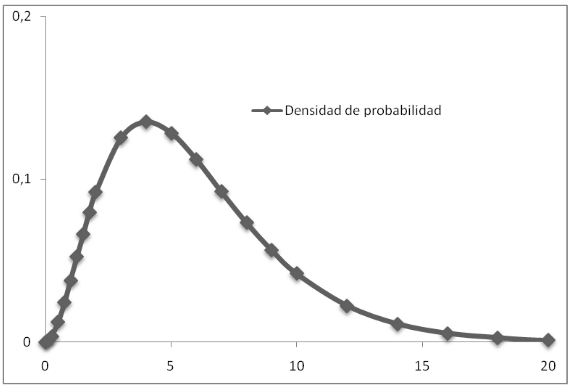
La probabilidad de que exista 1 exceso,

$$F(1) = P(X \leq 1) = \frac{1}{2^3 * 2} * \int_0^1 u^{3-1} * e^{-\frac{u}{2}} du =$$

Resolviendo la integral con ayuda del Derive, la probabilidad obtenida es,

$$= \frac{1}{16} * (16 - 26 * e^{-0,5}) = 0,01438$$

La representación gráfica y verificación con excel del cálculo:



Valores de x	Densidad de probabilidad	Probabilidad Acumulada	Probabilidad por la derecha
0	0	0	1
0,09	0,000483974	1,4684E-05	0,999985316
0,25	0,003447254	0,000296478	0,999703522
0,5	0,012168762	0,002161497	0,997838503
0,75	0,024162514	0,006652214	0,993347786
1	0,037908166	0,014387678	0,985612322
1,25	0,052271624	0,025656931	0,974343069
1,5	0,066426546	0,040505439	0,959494561
1,75	0,079789996	0,058803721	0,941196279
2	0,09196986	0,080301396	0,919698604
3	0,125510715	0,191153169	0,808846831
4	0,135335283	0,32332358	0,67667642
5	0,12825781	0,456186873	0,543813127
6	0,112020904	0,576809919	0,423190081
7	0,092479487	0,679152801	0,320847199
8	0,073262556	0,761896694	0,238103306
9	0,056239295	0,826421929	0,173578071
10	0,042112169	0,875347981	0,124652019
12	0,02230877	0,938031196	0,061968804
14	0,011170554	0,970363836	0,029636164
16	0,005367402	0,986246032	0,013753968
18	0,002499049	0,993767805	0,006232195
20	0,001134998	0,997230604	0,002769396
25	0,000145572	0,999658545	0,000341455
30	1,7207E-05	0,999960692	3,93084E-05
40	2,06115E-07	0,999999544	4,55515E-07
90	1,44915E-17	1	0

Ejemplo 6.

Si se sabe que el tiempo de sobrevivencia de ratas expuestas a un determinado tóxico es una variable aleatoria que sigue una distribución Gamma (5, 10), ¿cuál es la probabilidad de que una rata no supere las 60 semanas de vida?

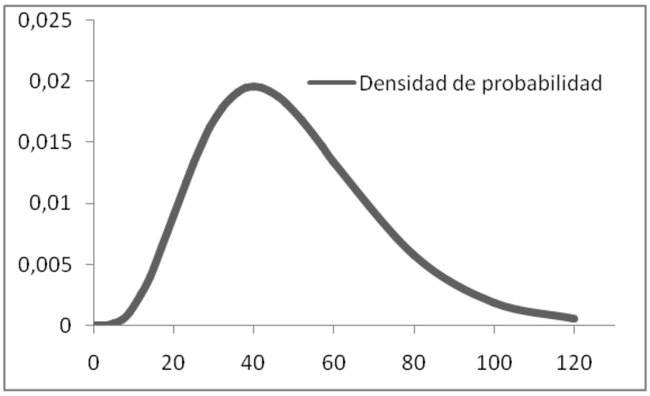
$$P(X < 60) = \frac{1}{10^5(5-1)!} * \int_0^{60} x^{5-1} * e^{-\frac{x}{10}} * dx$$

Resolviendo en R,

> pgamma(60, 5, scale = 10, lower.tail = T)

[1] 0.7149435

Su representación gráfica en excel



Fuentes:

http://gredos.usal.es/jspui/handle/10366/83459
http://www.coini.com.ar/COINI%202009/COINI%202008/Trabajos/TC3.pdf
http://departamentos.unican.es/macc/asignaturas/informatica/estadistica/pdf/Cap4.pdf
http://unbarquero.blogspot.com/2009/06/r-distribucion-gamma.html
http://www.suratep.com/articulos/27/accidentalidad_vial.pdf
http://upcommons.upc.edu/pfc/bitstream/2099.1/8321/1/Memoria%20PFC.pdf
http://www.aiaccess.net/English/Glossaries/GlosMod/e_gm_gamma_distri.htm
http://upcommons.upc.edu/revistes/bitstream/2099/5616/1/Article03a.pdf
http://office.microsoft.com/es-mx/excel-help/distr-gamma-HP005209101.aspx