Distribución Gamma

Kevin Piterman

19 de Mayo, 2020

¿Qué es la distribución Gamma y para qué sirve?

Es una distribución de probabilidades que nos permitirá **predecir el tiempo de espera hasta que ocurran ciertos eventos futuros**.

Esto es muy parecido a lo que vimos con la **Exponencial**: el tiempo de espera hasta que ocurra **el primer evento**.

En este caso, la distribución **Gamma** nos dice **el tiempo de espera hasta que ocurra el n-ésimo evento**.

Recapitulando: supongamos que que tenemos un proceso de Poisson $(Y_t)_{t>0}$ de parámetro λ. Si X es la variable que mide el tiempo de espera hasta que ocurre el primer evento, entonces sabemos que X es exponencial de parámetro λ. Concretamente,

$$X = \min\{t \,:\, Y_t > 0\} \quad \Longrightarrow \quad X \sim \mathcal{E}(\lambda).$$

Si X_n es la variable aleatoria que mide la cantidad de tiempo transcurrido hasta que ocurren n eventos en el proceso de Poisson, entonces X_n tiene distribución Gamma con parámetros n y λ. Concretamente,

$$X_n = \min\{t : Y_t \geq n\} \quad \Longrightarrow \quad X_n \sim \Gamma(n, \lambda).$$

Definición. Una variale aleatoria continua X tiene distribución Gamma con parámetros (α, λ), $\alpha, \lambda > 0$, si su densidad es

$$f(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} I_{(0,+\infty)}(x),$$

donde $\Gamma(\alpha)$ es la función gamma y está definida por:

$$\Gamma(\alpha) = \int_0^{+\infty} e^{-y} y^{\alpha-1} dy.$$

Notación: $X \sim \Gamma(\alpha, \lambda)$.

$$\alpha = \text{ cantidad de eventos que estoy esperando que ocurran},$$

$$\lambda = \text{cada cuánto ocurren los eventos}.$$

Propiedades.

1. (Escalar) $X \sim \Gamma(\alpha, \lambda)$ y $c > 0$ entonces $cX \sim \Gamma(\alpha, \frac{\lambda}{c})$.
2. $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$.
3. $\Gamma(n) = (n - 1)!$.
4. $\mathbb{E}(\Gamma(\alpha, \lambda)) = \frac{\alpha}{\lambda}$.
5. $V(\Gamma(\alpha, \lambda)) = \frac{\alpha}{\lambda^2}$.

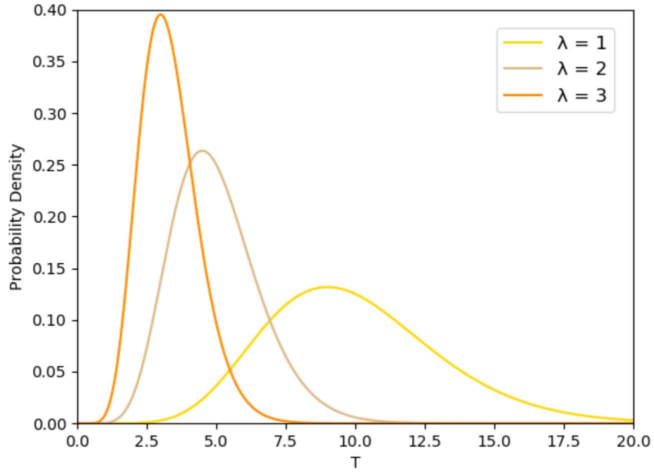

Figure 1: Función de densidad de Gamma con $\alpha = 10$.

Notar que de la definición es casi inmediata la siguiente propiedad:

Propiedad. (Suma) Si $X_1 \sim \Gamma(\alpha, \lambda)$ y $X_2 \sim \Gamma(\beta, \lambda)$ son independientes, entonces $X_1 + X_2 \sim \Gamma(\alpha + \beta, \lambda)$.

La demostración de esta propiedad se puede hacer con funciones generadoras de momentos (más adelante).

Desfío: intentar probarlo a mano (calculando su densidad como puedan).

Ejemplo 1. Estamos en un supermercado y tenemos que hacer la cola para pagar. Hay dos cajeros disponibles, A y B, que tienen 3 y 4 personas respectivamente. La cantidad de personas que puede despechar el cajero A sigue una distribución de Poisson con media 8 minutos. El cajero B despacha a una persona en 4 minutos con probabilidad 0.8 y en 10 minutos con probabilidad 0.2. El tiempo que tardan con cada persona es independiente, y los cajeros son independientes entre sí.

Elegimos al azar uno de los cajeros y hacemos la cola en ese. Calcular la probabilidad de haber elegido al cajero que vaya a tardar más.

Solución. Supongamos que A es el evento elegir al cajero A. Ídem con B. Queremos calcular entonces la siguiente probabilidad.

$$p = P(\text{elegir A y tarda más que B}) + P(\text{elegir B y tarda más que A})$$

Sea T_A el tiempo que tarda el cajero A en despechar a las 3 personas que tiene en la cola, y sea T_B el tiempo que tarda B en despachar a las 4 personas que tiene en la cola. Haciendo Probabilidad Total, vemos que p es la siguiente probabilidad:

$$p = P(T_A > T_B|A)P(A) + P(T_B > T_A|B)P(B) = P(T_A > T_B|A)\frac{1}{2} + P(T_B > T_A|B)\frac{1}{2}.$$

Como los cajeros atienden a los clientes de manera independiente, las probabilidades son independientes de qué cajero hayamos elegido, y por lo tanto

$$P(T_A > T_B|A) = P(T_A > T_B),$$
$$P(T_B > T_A|B) = P(T_B > T_A).$$

Para calcular $P(T_A > T_B)$ utilizamos probabilidad total sobre los valores que puede tomar T_B dado que T_B es una variable aleatoria discreta y además es independiente de T_A. Por otro lado, sabemos que $T_A \sim \Gamma(3, 8)$.

Para calcular la distribución de T_B, nos fijamos qué valores puede tomar:

$$R_{T_B} = \{4 * 4, 3 * 4 + 10, 2 * 4 + 2 * 10, 4 + 3 * 10, 4 * 10\} = \{16, 22, 28, 34, 40\}.$$

Sean $x_0 = 16$, $x_1 = 22$, $x_2 = 28$, $x_3 = 34$, $x_4 = 40$. Entonces el comportamiento de T_B es como el de una binomial, donde $P(T_B = x_i)$ es la probabilidad de haber atendido a i de los 4 clientes en 10 minutos. Es decir, es la probabilidad de que una binomial $\mathcal{B}i(4, 0.2)$ tome el valor i.

$$P(T_B = x_i) = \binom{4}{i} 0.2^i (0.8)^{4-i}.$$

Con esto podemos calcular las probabilidades anteriores:

$$\begin{aligned} P(T_A > T_B) &= \sum_i P(T_A > T_B|T_B = x_i)P(T_B = x_i) = \sum_i P(T_A > x_i|T_B = x_i)P(T_B = x_i) \\ &= \sum_i P(T_A > x_i)P(T_B = x_i). \end{aligned}$$

$$\begin{aligned} P(T_A < T_B) &= \sum_i P(T_A < T_B|T_B = x_i)P(T_B = x_i) = \sum_i P(T_A < x_i|T_B = x_i)P(T_B = x_i) \\ &= \sum_i P(T_A < x_i)P(T_B = x_i). \end{aligned}$$

Por lo tanto, sumando todo,

$$p = \frac{1}{2}\sum_i (P(T_A > x_i) + P(T_A < x_i))P(T_B = x_i) = \frac{1}{2}\sum_i 1P(T_B = x_i) = \frac{1}{2}.$$

Al final no necesitamos conocer la distribución de T_A ni de T_B. ¿Por qué sucedió esto? Una explicación sencilla es la siguiente. Seguro que un cajero va a tardar menos que otro porque uno de ellos sigue una distribución continua. Como elegimos al cajero al azar, tenemos $\frac{1}{2}$ de chances de haber elegido bien.

Este probabilidad no hubiera sido $\frac{1}{2}$ si los cajeros no hubieran sido elegidos al azar (por ejemplo, si tenemos preferencia por uno de los cajeros).

Ejemplo 2. La cantidad de lluvia que cae anualmente en Mendoza se modela con una variable aleatoria cuya función de distribución acumulada es

$$F(t) = 1 - e^{-0.2t} - 0.2te^{-0.2t} - 0.02t^2e^{-0.2t}, \quad t > 0$$

Determinar la media y la varianza de la caída de lluvia en Mendoza.

Solución. Podemos inferir que la variable aleatoria es continua pues su función de distribución acumulada es diferenciable. Para calcular su densidad, derivamos F. Hacemos todo para $t > 0$.

Sacamos primero factor común:

$$F(t) = 1 - e^{-0.2t}(1 + 0.2t + 0.02t^2)$$

Sea $\lambda = 0.2$.

$$F(t) = 1 - e^{-\lambda t}(1 + \lambda t + \frac{\lambda}{10}t^2)$$

$$F'(t) = \lambda e^{-\lambda t}(1 + \lambda t + \frac{\lambda}{10}t^2) - e^{-\lambda t}(\lambda + \frac{\lambda}{5}t)$$

$$F'(t) = \lambda e^{-\lambda t}(1 + \lambda t + \frac{\lambda}{10}t^2 - 1 - \frac{1}{5}t)$$

$$F'(t) = \lambda e^{-\lambda t}((\lambda - \frac{1}{5})t + \frac{\lambda}{10}t^2)$$

$$F'(t) = \lambda e^{-\lambda t}(\frac{\lambda}{10}t^2)$$

$$F'(t) = \lambda e^{-\lambda t}\frac{\lambda^2}{2}t^2$$

$$F'(t) = \frac{\lambda e^{-\lambda t}\lambda^2 t^2}{2}$$

$$F'(t) = \frac{\lambda e^{-\lambda t}(\lambda t)^{3-1}}{\Gamma(3)}$$

Por lo tanto, como la densidad de X coincide con la densidad de una Gamma de parámetros $(3, 0.2)$, concluimos que $X \sim \Gamma(3, 0.2)$.

Luego $\mathbb{E}(X) = \frac{\alpha}{\lambda} = 3 * 5 = 15$ y $V(X) = \frac{\alpha}{\lambda^2} = 3 * 25 = 75$.

Ejemplo 3. Simular una variable aleatoria $W \sim \Gamma(n, \lambda)$.

Solución. Sabemos que $W = \sum_{i=1}^{n} W_i$ con $W_i \sim \mathcal{E}(\lambda)$ independientes.

¿Cómo generamos una exponencial?

Si no nos acordamos cómo se hacía, lo hacemos a mano desde 0. Agarramos una uniforme $U \sim (0,1)$, y buscamos una función g de manera que $g(U) \sim \mathcal{E}(\lambda)$. Notemos que $g : (0,1) \to (0, +\infty)$ dado que la uniforme toma valores en el $(0,1)$ y la exponencial que queremos que sea $g(U)$ debe tomar valores en el intervalo $(0, +\infty)$.

Para calcular g, igualamos la función de distribución acumulada de una exponencial a la función de distribución acumulada de la variable $g(U)$. A partir de esta ecuación vamos a tratar de despejar a la función g. Para esto vamos a suponer que g es creciente. Sea $t \geq 0$. Entonces:

$$1 - e^{-\lambda t} = F_{\mathcal{E}(\lambda)}(t) = F_{g(U)}(t) = P(g(U) \leq t) = P(U \leq g^{-1}(t)) = F_U(g^{-1}(t)) = g^{-1}(t).$$

Hacemos un cambio de variable para que aparezca g y no g^{-1}. Sea $u = g^{-1}(t)$, por lo que $g(u) = t$. Con esto nos queda que:

$$1 - e^{-\lambda g(u)} = u$$

y despejando,

$$1 - u = e^{-\lambda g(u)}$$

$$\ln(1-u) = -\lambda g(u)$$

$$-\frac{1}{\lambda} \ln(1-u) = g(u).$$

Por lo tanto, $-\frac{1}{\lambda} \ln(1-U) \sim \mathcal{E}(\lambda)$.

Para terminar, generamos n uniformes independientes $U_i \sim (0,1)$ y consideramos $W_i = -\frac{1}{\lambda} \ln(1-U_i)$. Luego

$$W = \sum_i -\frac{1}{\lambda} \ln(1-U_i) = -\frac{1}{\lambda} \sum_i \ln(1-U_i) = -\frac{1}{\lambda} \ln\left((1-U_1)\dots(1-U_n)\right).$$