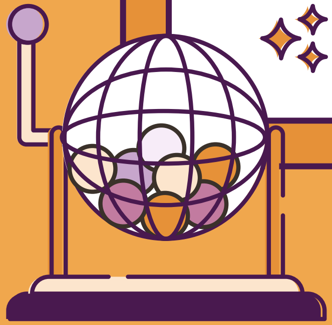
Matemáticas para la ciencia de Datos

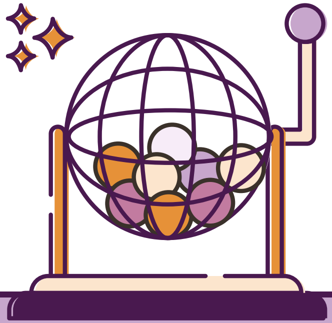
Distribución Gamma

Leonel García
Eduardo Hernández
Luis Izquierdo
Salvador Medina
Mauricio Tellez



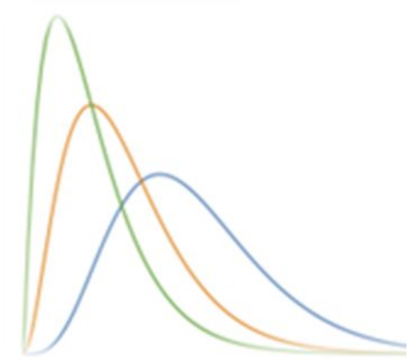
1.

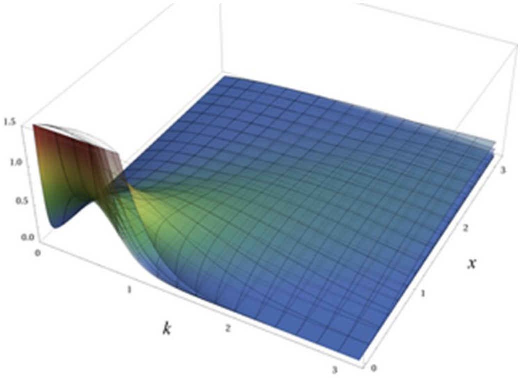
Introducción



Introducción

La **distribución gamma** es una de las distribuciones de probabilidad continua fundamentales en estadística, caracterizada por modelar la acumulación de tiempos o cantidades. A lo largo de las décadas, esta distribución ha sido esencial para describir eventos que se desarrollan en el tiempo o en procesos donde se acumulan cantidades hasta cierto umbral.





Fuente: Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L

Introducción

La distribución gamma se caracteriza por dos parámetros principales: el **parámetro de forma** α y el **parámetro de escala** β. La variación de estos parámetros afecta la forma de la distribución:

•**α** controla la "asimetría" y la concentración de los valores, con valores más altos de α resultando en una distribución menos sesgada.

•**β** influye en la dispersión o varianza, afectando cómo se distribuyen los valores en torno a la media.

La importancia de la distribución gamma se debe a su habilidad para modelar tiempos de espera acumulados en procesos aleatorios, como el tiempo hasta el **α** -ésimo evento en un proceso de Poisson. Su función de densidad de probabilidad es:

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

Fuente: Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L

2.

Estado del arte

Estado del arte
"Orígenes de la Función Gamma"

La historia de la distribución gamma comienza indirectamente con el desarrollo de la función. Fue introducida en el siglo XVIII:

•**Leonhard Euler (1729)**: Euler propuso la función gamma como una extensión del concepto de factorial para números reales y complejos. Su trabajo fue fundamental porque los factoriales sólo estaban definidos para números enteros positivos, y la función gamma extendió esta definición, facilitando su uso en análisis matemático y probabilidad.

•**Legendre y Gauss (Siglo XVIII-XIX)**: Continuaron trabajando con la función gamma, lo que impulsó su uso en diferentes contextos matemáticos, aunque aún no se vinculaba directamente con distribuciones de probabilidad.

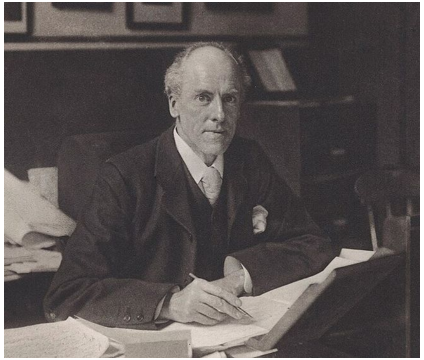
Fuente: Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L

Estado del arte
“Primeros desarrollos en Probabilidad y Estadística”

La teoría de probabilidades experimentó un desarrollo importante en los siglos XIX y XX, y el papel de la función gamma en probabilidades comenzó a emerger en ese contexto:

•**Karl Pearson (1900s)**: Pearson fue uno de los primeros en estudiar distribuciones continuas y su relación con la función gamma. Desarrolló el sistema de distribuciones de Pearson, del cual la distribución gamma es un caso especial.

•**Fisher y Tippett (1920s)**: Estos estadísticos británicos exploraron las distribuciones de probabilidades en relación con tiempos de vida y eventos extremos, lo que sentó las bases para aplicaciones prácticas de la distribución gamma en tiempos de espera y tiempos hasta eventos.



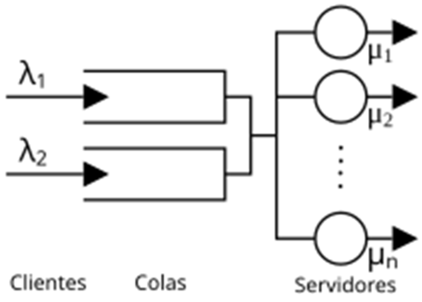
Fuente: Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L

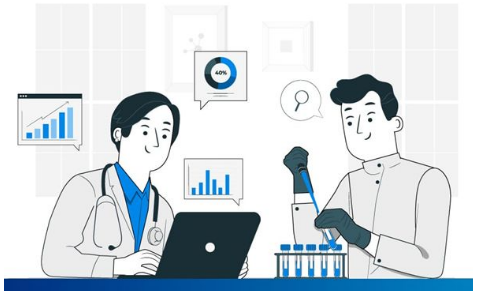
Estado del arte
“Formalización de la Distribución Gamma”

Durante el siglo XX, matemáticos y estadísticos formalizaron la distribución gamma como una herramienta estadística para describir procesos estocásticos:

•**Procesos de Poisson y Distribución Gamma**: La distribución gamma comenzó a relacionarse explícitamente con el tiempo de espera hasta el n-ésimo evento en un proceso de Poisson. Esta relación cimentó su utilidad en modelar fenómenos de espera en los que los eventos ocurren de manera independiente y acumulativa.

•**Aplicaciones en Ciencias e Ingeniería (1950s-1970s)**: A mediados del siglo XX, la distribución gamma se volvió popular en diversas aplicaciones, como en el modelado de vida útil de dispositivos (ingeniería de confiabilidad), flujos de tráfico, tiempos de respuesta de sistemas y procesos de desintegración radiactiva.





Fuente: Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L

Estado del arte
"Expansión y aplicaciones actuales"

Los usos de la distribución gamma se han expandido con el desarrollo de nuevas teorías y tecnologías:

•**Teoría de Procesos Estocásticos y de Colas**: Es fundamental para modelar los tiempos de espera en procesos de colas y en redes de telecomunicaciones.

•**Bioestadística y Modelos de Supervivencia**: Se usa en estudios de supervivencia para modelar tiempos de vida de organismos o tiempos hasta que ocurre un evento biológico específico.

•**Aplicaciones en Finanzas**: En la modelación de riesgos y tiempos de espera en transacciones financieras, la distribución gamma ha encontrado aplicación en la modelación de volatilidades y tiempos hasta eventos financieros importantes.

•**Machine Learning y Análisis de Datos**: Con la creciente disponibilidad de datos y algoritmos, la distribución gamma se emplea en modelos bayesianos y de aprendizaje automático.

Fuente: Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L

3.

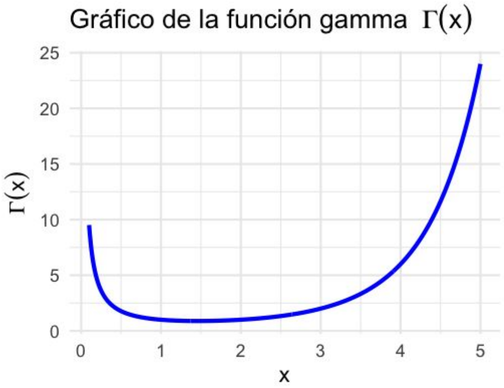
Desarrollo

Distribución gamma
Función gamma

En cursos de cálculo avanzado se habla sobre la existencia de la integral para $\alpha > 0$:

$$\int_0^\infty y^{\alpha-1} e^{-y} dy$$

Igual se demuestra que su valor es positivo. Esta integral se llama función gamma de α y se denota con $\Gamma(\alpha)$.



Fuente: Hogg, Craig

Distribución gamma
Función gamma

Una de las propiedades de esta función es:

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

Para $\alpha > 1$ se puede demostrar integrando por partes que:

$$\Gamma(\alpha) = (\alpha - 1) \int_0^\infty y^{\alpha-2} e^{-y} dy = (\alpha - 1)\Gamma(\alpha - 1)$$

De manera particular, si α es entero positivo:

$$\Gamma(\alpha) = (\alpha - 1)(\alpha - 2) \dots (2)(1) = (\alpha - 1)!$$

Fuente: Hogg, Craig

Distribución gamma
Función gamma

Note que al realizar en $\Gamma(\alpha)$ el cambio de variable $y = x/\beta$ para $\beta > 0$ se obtiene:

$$\Gamma(\alpha) = \int_0^\infty \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\frac{x}{\beta}} \left(\frac{1}{\beta}\right) dx$$

Y de manera equivalente:

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

Como $\alpha, \beta, \Gamma(\alpha) > 0$ observamos que:

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

es una función de densidad de una variable aleatoria continua en $(0, \infty)$. A la variable aleatoria con función de densidad $f(x)$ se dirá que tiene distribución gamma con parámetros α y β.

Fuente: Hogg, Craig

Distribución gamma
Definición

Es una distribución continua con soporte $(0, \infty)$ que tiene dos parámetros. Comunmente existen dos versiones:

Parámetros de forma *k* y escala θ

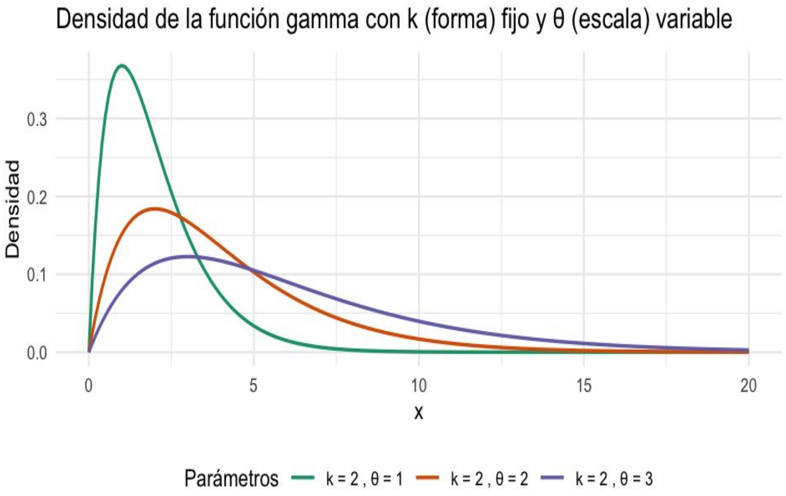
$$f(x; k, \theta) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}} I_{(0,\infty)}(x)$$

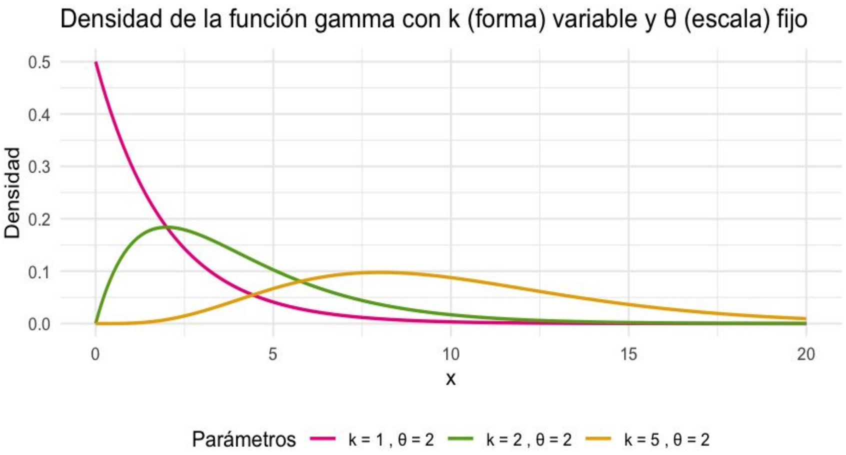
Parámetros de forma α y tasa β (recíproco de escala)

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

Fuente: [Hogg, Craig]

Distribución gamma
Definición





Fuente: Hogg, Craig

Distribución gamma
Observaciones

- En procesos estocásticos y análisis de supervivencia se utiliza para modelar tiempos de espera.
- Al no tener una forma cerrada, su distribución acumulada tampoco la tiene. Simplemente se deja expresada.
- Al ser monomodal y tener soporte en los números positivos, es muy utilizada para modelar diversos fenómenos.

Fuente: Hogg, Craig

Distribución gamma
Propiedades

Momentos: Si X es una variable aleatoria gamma con parámetros k y θ, tiene los siguientes momentos:

$$E(X) = k\theta$$
$$E\left(X^2\right) = k\theta^2 + k^2\theta^2$$

La función generadora de momentos es:

$$M(t) = (1 - \theta t)^{-k}, t < 1/\theta$$

Distribución gamma
Propiedades

- Si $X_1, \dots, X_n$ son v.a.s independientes con $X_i \sim Gamma(k_i, \theta)$, entonces $S = X_1 + \cdots + X_n \sim Gamma(\sum k_i, \theta)$.
- Para un conjunto de *n* variables aleatorias $X_1, \dots, X_n$ exponenciales iid de parámetro λ, la variable aleatoria $S = \sum X_i$ tendrá distribución Gamma con parámetros de forma *n* y de tasa λ.
- Si X es una variable aleatoria Gamma(1, λ) (de forma y tasa), entonces X tiene distribución exponencial de parámetro λ.
- Una variable aleatoria *X* con distribución gamma y parámetro de forma *k* grande, converge en distribución a una variable aleatoria normal de media $k\theta$ y varianza $k\theta^2$. Esto es por el TLC.

Fuente: Kaas, et al.

Distribución gamma
Gamma trasladada

- La distribución de las reclamaciones totales *S* usualmente es sesgada a la derecha ($\gamma > 0$), tiene sopoerte no negativo y es unimodal.

- La distribución gamma se utiliza para aproximar la distribución de las reclamaciones totales 'moviendo' a la usual distribución gamma, utilizando $S \approx X + x_0$ donde $X \sim Gamma\ (\alpha, \beta)$ y x_0 es una constante.

- En este sentido tendremos que:

$$F_S(s) \approx P(X \leq s - x_0)$$

- Para estimar los parámetros α, β y x_0 se igualan los momentos de $X+x_0$ con los momentos de *S*.

Fuente: [Kaas, et al.]

Distribución gamma
Gamma trasladada

[Ejemplo 2.5.5, Kaas] El monto total S de las reclamaciones tiene valor esperado estimado 10,000, desviación estándar 1,000 y sesgo $\gamma = 1$. Se puede demostrar que:

$$\hat{\alpha} = \frac{4}{\gamma^2}, \hat{\beta} = \frac{2}{\gamma\sigma}, x_0 = \mu - \frac{2\sigma}{\gamma}$$

En este sentido tenemos:

$$\hat{\alpha} = 4, \hat{\beta} = 0.002, x_0 = 8{,}000$$

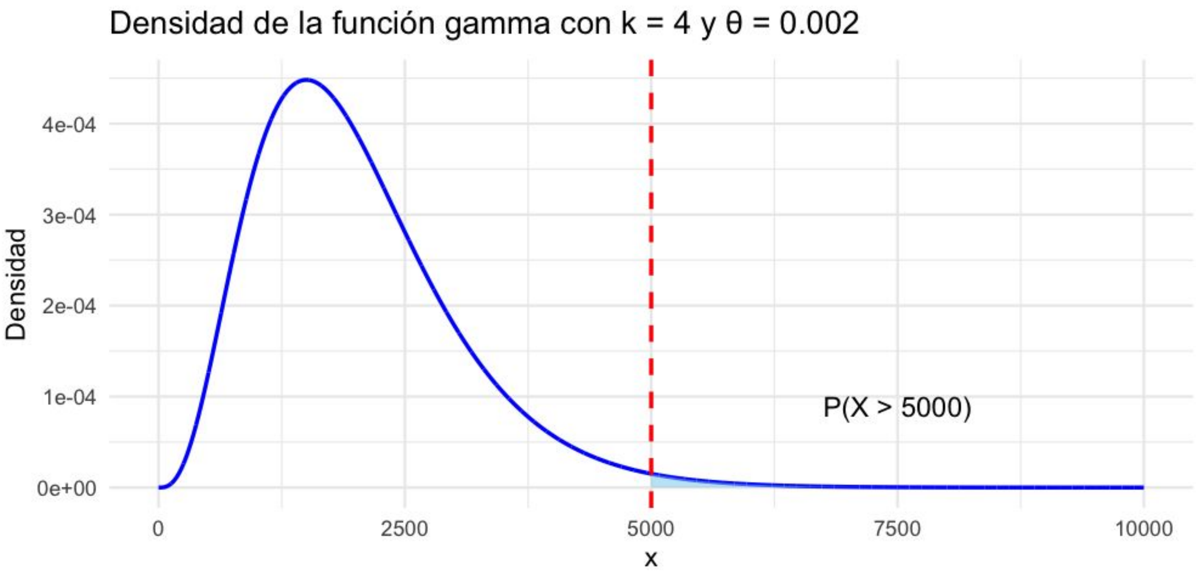
Si queremos calcular, por ejemplo, la probabilidad de que el monto total de las reclamaciones supere $13,000 calculamos:

$$P(S > 13{,}000) = 1 - P(S \leq 13{,}000) \approx 1 - P(X \leq 5{,}000) = 0.01$$

Donde $X \sim Gamma\ (4{,}0.002)$.

Fuente: Kaas, et al.

Distribución gamma
Gamma trasladada



Fuente: Kaas, et al.

4.

Casos de uso

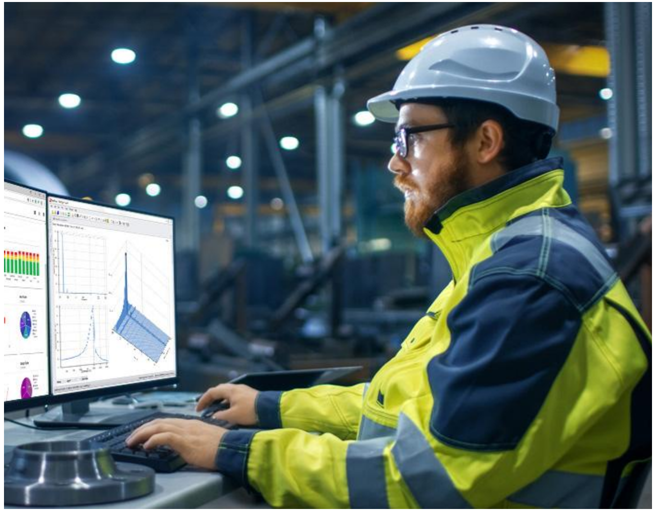
Mantenimiento de equipos industriales

En su trabajo *Cálculo de la mantenibilidad usando la distribución gamma,* Benítez-Montalvo et al. (2018), se proponen evaluar la mantenibilidad de equipos industriales, mediante el análisis de datos recolectados entre 2009 y 2015 de un sistema de gestión de mantenimiento.

Busca identificar la distribución adecuada para modelar los tiempos de reparación, contrastando distintas distribuciones como la normal, exponencial y Weibull con respecto a la distribución Gamma.

Fuente: Benítez-Montalvo, R. I., Díaz-Concepción, A., Rodríguez Piñeiro, A. J., Villar Ledo, L., & Rodríguez Perez, H. (2018).

Mantenimiento de equipos industriales



El estudio concluye que la distribución Gamma es eficaz para el cálculo de la mantenibilidad, al permitir una estimación de los tiempos de reparación con un error más bajo. Añade valor a la empresa propietaria de los equipos, cuando se instrumentaliza en un análisis de ahorro potencial, basado en las estimaciones del modelo.

Fuente: Benítez-Montalvo, R. I., Díaz-Concepción, A., Rodríguez Piñeiro, A. J., Villar Ledo, L., & Rodríguez Perez, H. (2018).

Distribución de la renta

El objeto de estudio *La distribución Gamma como modelo para analizar la distribución de la Renta: una aplicación a la E.P.F. 1990-91* (Lechuga, M. L., 1998) es la desigualdad en la distribución de ingresos entre las Comunidades Autónomas en España.

Usa la distribución Gamma para modelar los ingresos y calcular índices de desigualdad como el Índice de Gini, Atkinson y Entropía Generalizada. También emplea métodos como el Máximo de Verosimilitud para estimar los parámetros de la distribución Gamma, y se validan los resultados utilizando el test de Kolmogorov-Smirnov para asegurar que los datos se ajusten correctamente al modelo.

Fuente: Lechuga, M. L. (1998).

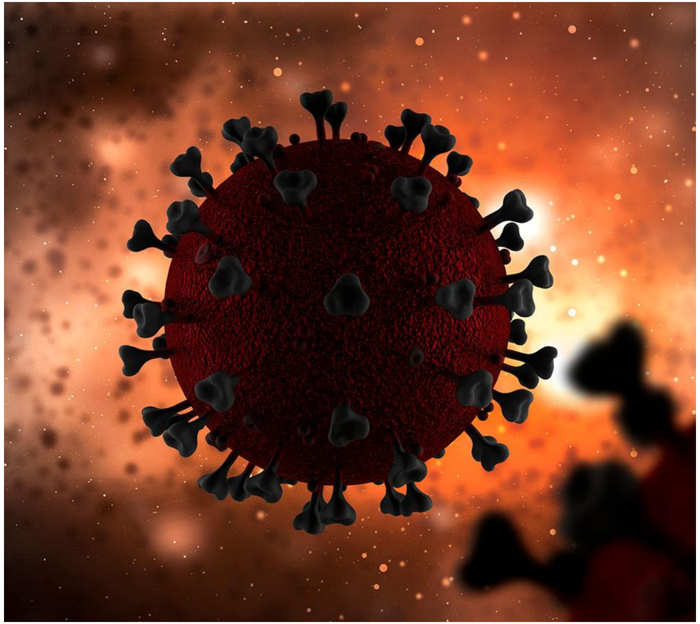
Distribución de la renta

Los resultados arrojaron que, para el periodo de análisis, comunidades como Asturias y Navarra presentaron menor desigualdad, mientras que regiones como Murcia, Andalucía y Extremadura presentan una mayor desigualdad en sus ingresos. Lo cual tiene implicaciones muy relevantes en el diseño de políticas públicas.

Fuente: Lechuga, M. L. (1998).

Epidemiología

Durante las fases iniciales de la pandemia por COVID-19, Estrada-Alvarez, J. M. et. al. (2023) se propusieron conocer la dinámica de transmisión del virus en la región de Pereira, Colombia.



En su trabajo *Estimación del intervalo serial y número reproductivo básico para los casos importados de COVID-19*, recopilaron datos de campo mediante entrevistas epidemiológicas a 12 casos confirmados y sus contactos. Ajustaron modelos utilizando la distribución Gamma para estimar el intervalo serial y el R0, aplicando el método de máxima verosimilitud.

Fuente: Estrada-Alvarez, J. M., Ospina-Ramírez, J. J., Hincapié-Acuña, M., & Gómez-González, M. D. P. (2023).

Epidemiología

El estudio encontró que el intervalo serial promedio fue de 3.8 días, más corto que el período de incubación, sugiriendo una transmisión presintomática significativa. La importancia de esta investigación radica en su capacidad para ofrecer información importante en el diseño de estrategias de salud efectivas, como el rastreo temprano de contactos y el aislamiento preventivo, lo cual, a su vez, es fundamental en los esfuerzos de prevención de contagios.

Fuente: Estrada-Alvarez, J. M., Ospina-Ramírez, J. J., Hincapié-Acuña, M., & Gómez-González, M. D. P. (2023).

Agricultura

Para enfrentar la incertidumbre sobre las precipitaciones en zonas agrícolas de Zacatecas, Magallanes-Quintanar y Valdez-Cepeda (2011) realizaron un estudio utilizando la distribución Gamma para modelar y predecir las probabilidades de lluvia.

Emplearon datos históricos de 32 años de precipitación registrados en una estación meteorológica. Utilizando la distribución Gamma y el método de estimadores de momentos, se ajustaron modelos para calcular las probabilidades de ocurrencia de diferentes niveles de lluvia. Se compararon los resultados con la distribución normal para validar el mejor ajuste, utilizando la prueba de bondad de ajuste de Smirnov.

Fuente: Magallanes-Quintanar, R., & Valdez-Cepeda, R. D. (2011)

Agricultura

El análisis concluyó que la distribución Gamma pudo capturar mejor la variabilidad y asimetría de los datos de precipitación. Esto permite predecir con mayor exactitud las probabilidades de lluvia, y proporciona a los agricultores una herramienta que mejora la planificación de cultivos y el uso eficiente del agua.

Fuente: Magallanes-Quintanar, R., & Valdez-Cepeda, R. D. (2011)

5.

Ejemplos

Ejemplo 1 (ciencias de la tierra) : Distribución Gamma

- La magnitud de temblores registrados en una región de América del Norte puede modelarse como si tuviera una distribución exponencial con media 2.4, según se mide en la escala de Richter. Encuentre la probabilidad de que un temblor que ocurra en esta región

a) sea mayor que 3.0 en la escala de Richter.

b) caiga entre 2.0 y 3.0 en la escala de Richter.

Nota: Una distribución exponencial con parámetro de tasa λ es un caso especial de la distribución Gamma, donde el parámetro de forma $\alpha=1$ y el parámetro de escala $\theta=1/\lambda$

Fuente: Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L

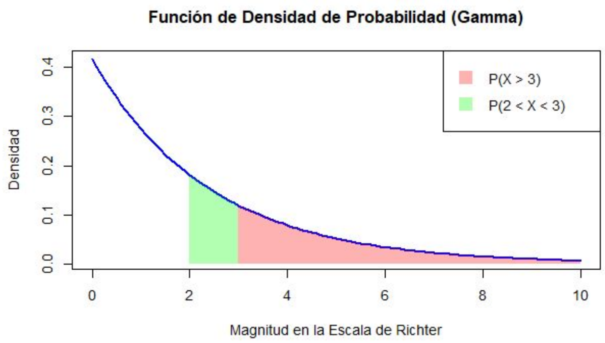
Cálculo en R

```r
# Parámetros
alpha <- 1  # Parámetro de forma
beta <- 2.4  # Parámetro de escala para la gamma

# Inciso (a): Probabilidad de que X > 3
p_a <- 1 - pgamma(3, shape = alpha, scale = beta)
cat("Probabilidad de que X > 3:", p_a, "\n")

# Inciso (b): Probabilidad de que 2 < X < 3
p_b <- pgamma(3, shape = alpha, scale = beta) - pgamma(2, shape = alpha, scale = beta)
cat("Probabilidad de que 2 < X < 3:", p_b, "\n")
```

```
Probabilidad de que X > 3: 0.2865048
Probabilidad de que 2 < X < 3: 0.1480934
```



- La probabilidad de que la magnitud sea mayor que 3.0 es aproximadamente 0.2865.
- La probabilidad de que la magnitud esté entre 2.0 y 3.0 es aproximadamente 0.1487.

Fuente: Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L

Cálculo Manual

- Puesto que la media de la distribución exponencial es 2.4, podemos establecer el parámetro de escala θ=2.4, y el parámetro de forma α=1 dado que estamos tratando con una distribución exponencial.

- La función acumulada para la distribución Gamma es:

$$F(x;\alpha,\theta)=\int_0^x \frac{t^{\alpha-1}e^{-\left(\frac{t}{\theta}\right)}}{\theta^{\alpha}\Gamma(\alpha)}dt$$

- Y cuando α=1, entonces, estamos tratando con una función exponencial, donde $\Gamma(1)=1$ y se simplifica a:

-> $F(x;\alpha,\theta)=\int_0^x \frac{t^{\alpha-1}e^{-\left(\frac{t}{\theta}\right)}}{\theta^{\alpha}\Gamma(\alpha)}dt=\int_0^x \frac{t^{1-1}e^{-\left(\frac{t}{\theta}\right)}}{\theta^{1}\Gamma(1)}dt=\int_0^x \frac{e^{-\left(\frac{t}{\theta}\right)}}{\theta}dt$

Fuente: Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L

a) Encuentre la probabilidad de que un temblor que ocurra en esta región sea mayor que 3.0 en la escala de Richter.

- Queremos calcular $P(X > 3)$, que es equivalente a $1 - F(3; \alpha, \theta)$, con nuestros parámetros, α=1 y θ=2.4

Calculamos la CDF de la distribución gamma en $\mathrm{X} = 3$:

$$\mathrm{F}(3; \alpha=1, \theta=2.4) = \int_0^3 \frac{t^{1-1} e^{-\left(\frac{t}{2.4}\right)}}{2.4^1 \Gamma(1)} dt = \int_0^3 \frac{e^{-\left(\frac{t}{2.4}\right)}}{2.4} dt$$

- Utilizando la calculadora de integrales Symmbolab, tenemos que:
- $\int_0^3 \frac{e^{-\left(\frac{t}{2.4}\right)}}{2.4} dt \approx 0.7135$, de ahí que:
- $\boxed{\mathrm{P}(\mathrm{X} > 3) = 1 - 0.7135 \approx 0.2865}$

- La probabilidad de que la magnitud sea mayor que 3.0 es aproximadamente 0.2865.

Fuente: Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L

b) Encuentre la probabilidad de que un temblor que ocurra en esta región esté entre 2.0 y 3.0 en la escala de Richter.

- Queremos calcular $P(2 < X < 3)$, que es equivalente a:
- $P(2 < X < 3) = F(3; \alpha = 1, \theta = 2.4) - F(2; \alpha = 1, \theta = 2.4)$

Calculamos la CDF de la distribución gamma en $X = 2$:

$$F(3; \alpha{=}1, \theta{=}2.4) = \int_0^2 \frac{t^{1-1} e^{-\left(\frac{t}{2.4}\right)}}{2.4^1 \Gamma(1)} dt = \int_0^2 \frac{e^{-\left(\frac{t}{2.4}\right)}}{2.4} dt$$

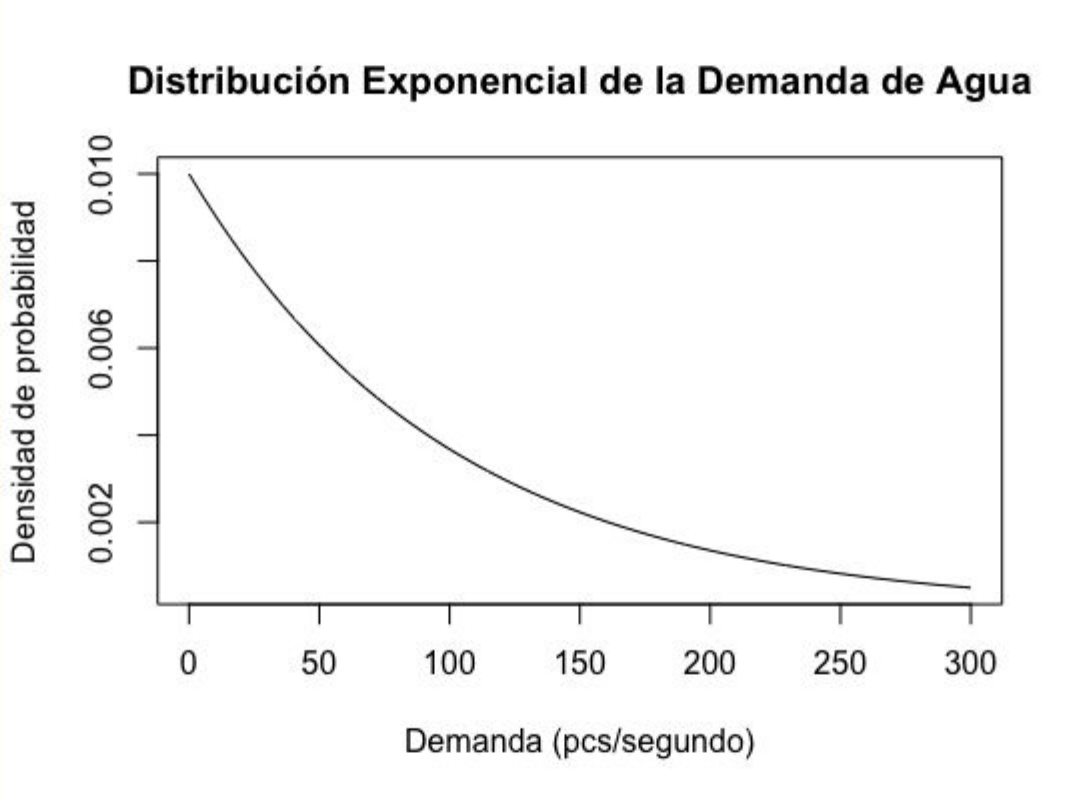
- Utilizando la calculadora de integrales Symmbolab, tenemos que:
- $\int_0^2 \frac{e^{-\left(\frac{t}{2.4}\right)}}{2.4} dt \approx 0.5648$, de ahí que:
- $\boxed{P(2 < X < 3) = 0.7135 - 0.5648 \approx 0.1487}$
- La probabilidad de que la magnitud esté entre 2.0 y 3.0 es aproximadamente 0.1487

Demanda de Agua

El operador de una estación de bombeo ha observado que la demanda de agua durante las primeras horas de la tarde tiene una distribución aproximadamente exponencial con media de 100 pes (pcs cúbicos por segundo).

- Encuentre la probabilidad de que la demanda sea mayor que 200 pcs durante las primeras horas de la tarde en un día seleccionado al azar.
- ¿Qué capacidad de bombeo de agua debe mantener la estación durante las primeras horas de la tarde para que la probabilidad de que la demanda sea mayor que la capacidad en un día seleccionado al azar sea de sólo .01?

Gráfico del modelo del problema



Solución

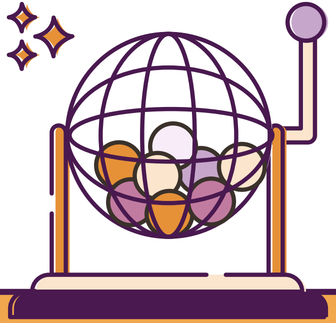
Datos proporcionados:
Media = 100 pcs

Probabilidad de que la demanda sea mayor que 200 pcs
$P(X) > 200$, donde X es la demanda de agua

Se calcula la tasa de ocurrencia (λ), que debe ser el inverso de la media:
$\lambda = 1/100 = 0.01$

Usando la Función de Distribución Acumulada:
$F(x) = 1 - e^{\wedge}(-\lambda x)$
$F(200) = 1\text{-}e^{\wedge}(-0.01*200) = 0.864$

Entonces, la probabilidad será:
$P(X) = 1-F(200) = 1 - .864 =$ **0.135** = **13.5%**



Solución

Capacidad de bombeo para una probabilidad de 0.01

$P(X > x) = 0.01$

Con la probabilidad sabemos que

$P(X) = 1 - F(X) = 0.01$

$F(X) = 0.99$

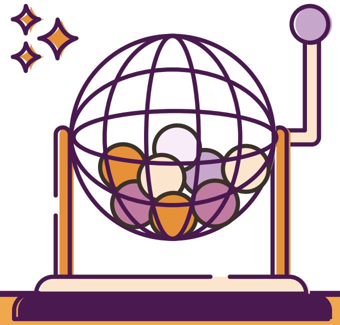
Sustituyendo y despejando en la Función de Distribución Acumulada

$1 - e^{(-0.01*x)} = 0.99$

$e^{(-0.01*x)} = 0.01$

$-0.01*x = \ln(0.01)$

$\mathbf{x = 460.517}$



6.

Conclusión

Conclusión

Distribución

Puede modelar variables aleatorias continuas positivas y con asimetría positiva

Flexibilida d

Debido a que puede adoptar una gran variedad de formas, se adapta a diferentes tipos de datos

Usos

- Tiempo entre eventos
- Duración de procesos
- Magnitudes

Referencias

- Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L. (s.f.). *Estadística matemática con aplicaciones* (7ª ed., p. 190, Ejercicio 4.88). Cengage Learning.
- Benítez-Montalvo, R. I., Díaz-Concepción, A., Rodríguez Piñeiro, A. J., Villar Ledo, L., & Rodríguez Perez, H. (2018). Cálculo de la mantenibilidad usando la distribución gamma. Ingeniería Mecánica, 21(1), 52-58.
- Lechuga, M. L. (1998). La distribución Gamma como modelo para analizar la distribución de la renta: una aplicación a la EPF 1990-91. Revista de Estudios Regionales, 1, 161-186.
- Estrada-Alvarez, J. M., Ospina-Ramírez, J. J., Hincapié-Acuña, M., & Gómez-González, M. D. P. (2023). Estimación del intervalo serial y número reproductivo básico para los casos importados de COVID-19. Revista de Salud Pública, 22, 194-197.
- Magallanes-Quintanar, R., & Valdez-Cepeda, R. D. (2011). Distribución Gamma xy cálculo de probabilidades de lluvia. Difu100ci@, Revista de difusión científica, ingeniería y tecnologías, 4(3), 67-71.
- Hogg, R., Craig, A. (1978). Introduction to Mathematical Statistics (4th ed). Macmillan Publishing.
- Kaas, R. (2008). Modern actuarial risk theory: Using R (2nd ed). Springer.