Transformation of a bivariate Γ-distribution

Paul Libbrecht, IUBH Advanced Statistics, CC-BY

2020-02-04

1 Statement

Let X and Y be independent random variables $\sim \Gamma(2,5)$; we call f_X and f_Y their density function. Compute the density function of the variable $T = X + Y$.

2 Straight of the theory

X and Y are random variables, that is they are $X : \mathcal{S} \longrightarrow \mathbb{R}^+$ and $Y : \mathcal{S} \longrightarrow \mathbb{R}^+$. They are Gamma variables of parameters $\alpha = 2$ and $\beta = 5$ (see [Ano21]), thus, they are distributed in such a way that given a set of real-numbers E:

$$P(Y \in E) = P(X \in E) = \int_{t \in E} \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t} dt = \int_{t \in E} \frac{5^2}{1} t e^{-5t} dt = \int_{t \in E} 25 t e^{-5t} dt$$

Let us call $f_X(t) = f_Y(t) = 25te^{-5t}$.

3 Transformations?

Given a multivariate continuous random variable $X : \mathcal{S} \to D$ (having values in any set) and a mapping $g : D \to E$ a mapping, the **transformed random variable** $g \circ X$, often written as $g(X)$ is defined by $g(X)(s) = g(X(s))$.

We assume $E, D \subset \mathbb{R}^n$. Then this is the same as saying: given a series of random variables $X_i : \mathcal{S} \longrightarrow \mathbb{R}$ where $0 \leq i \leq n$ and $(X_1, \ldots, X_n) \in D$ and a series of functions $g_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ where $(g_1(x_1, \ldots, x_n), \ldots, g_n(x_1, \ldots, x_n)) \in E$ when $(x_1, \ldots, x_n) \in E$.

Suppose that D and E are subsets of $\mathbb{R}^n$ and that g is a derivable tranformation (which means: $g_i(x_1, ..., x_n)$ is derivable, i.e. $\frac{\partial g_i}{\partial x_j}$ exists for each $0 \leq i, j \leq n$).

The transformation theorem, proved in [HMC20, 2.7], says that:

- $g(X)$ It is a continuous random variable.
- if $D, E \subset \mathbb{R}^n$ and X is a continuous random variable with pdf $f_X : D \to \mathbb{R}$ then the pdf of $g(X)$, noted $f_{g(X)}$, is equal to the following, $\forall \mathbf{a} \in E$:

$$f_{g(x)}(\mathbf{a}) = f_X(g^{-1}(\mathbf{a})) \cdot |J_{g^{-1}}(\mathbf{a})|$$

It is important to note that D and E are rarely equal to complete $\mathbb{R}^n$. E.g. It could be a subset of a plane where the lowest x depends on y.

4 Example

Coming back to our example random variable $T(s) = X(s) - Y(s)$ with $X \sim \Gamma(2,5)$ and $Y \sim \Gamma(2,5)$ we consider the random variable *couple* $C(s) = (X(s), Y(s))$ which gives a point for each possible random outcome (the points are in the upper-half quadrant $\{(x,y)|x > 0 \text{ and } y > 0\}$). To calculate the product we use the following transformation:

$$\begin{array}{rlll} g: & \mathbb{R}^+ \times \mathbb{R}^+ & \longrightarrow & E \\ & (x,y) & \mapsto & (a,b) = g(x,y) = (x-y, x+y) \end{array}$$

What is the destination subset $E \subset \mathbb{R} \times \mathbb{R}$ of the transformation g? It is the set of $(a,b) = g(x,y)$. That is, it is the set of (a,b) such that we can find $(x,y) \in D$ with $x - y = a$ and $x + y = b$. We can calculate this set by solving the equation to obtain x and y from a and b:

$$\begin{cases} a = x - y \\ b = x + y \end{cases} \underset{\text{to second}}{\overset{\text{add first line}}{\iff}} \begin{cases} a = x - y \\ a + b = 2x \end{cases}$$

$$\underset{\text{use expression of x}}{\overset{\substack{\text{express x from} \\ \text{second line}}}{\iff}} \begin{cases} x = \frac{1}{2}(a+b) \\ y = \frac{1}{2}(a+b) - a \end{cases} \underset{..}{\overset{\text{simplify}}{\iff}} \begin{cases} y = \frac{1}{2}(b-a) \\ x = \frac{1}{2}(a+b) \end{cases}$$

Thus, E is the set of pairs $(a,b) \in \mathbb{R} \times \mathbb{R}$ such that $b - a > 0$ and $b + a > 0$ which his the hatched region in Figure 2. From solving this equation, we can right away calculate the inverse transformation g^{-1}:

$$(x,y) = g^{-1}(a,b) = \left(\frac{1}{2}(a+b), \frac{1}{2}(b-a)\right)$$

We now apply the transformation theorem above with the transformation g so as to get the density of $g(C)$. This will give us the joint density of $(X-Y, X+Y)$ from which we shall be able to deduce the density of $X - Y$ as intended.

To apply the theorem, we need to calculate the Jacobian of g^{-1}:

$$J_{g^{-1}}(a,b) = \begin{vmatrix} \frac{\partial(\frac{1}{2}(b-a))}{\partial a} & \frac{\partial(\frac{1}{2}(b-a))}{\partial b} \\ \frac{\partial(\frac{1}{2}(b+a))}{\partial a} & \frac{\partial(\frac{1}{2}(b+a))}{\partial b} \end{vmatrix} = \frac{1}{2} \cdot \begin{vmatrix} -1 & 1 \\ 1 & 1 \end{vmatrix} = \frac{1}{2} \cdot (-2) = -1$$

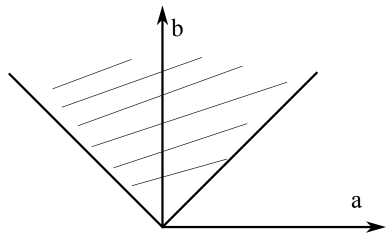

Figure 1: The set of (a, b) reached by the transformation g.

Let us remind that the density of the X and Y (both $\sim \Gamma(2,5)$) is given by the function $f(t) = \frac{25}{1} \cdot t^{2-1} \cdot e^{-5t} = 25te^{-5t}$ and they are independent thus the joint density of (X, Y) is:

$$f_{(X,Y)}(x, y) = 25xe^{-5x} \cdot 25ye^{-5y} = 625 \cdot x \cdot y \cdot e^{-5(x+y)}$$

The theorem thus gives us the joint density of $g(X, Y)$ is the following for $(a, b) \in E$ and is 0 otherwise:

$$f_{g(X,Y)}(a, b) = f_{(X,Y)}(g^{-1}(a, b)) \cdot |J_{g^{-1}}(a, b)| = \frac{625}{4} \cdot (b-a) \cdot (a+b) \cdot e^{-5 \cdot 2 \cdot b} = \frac{625}{4}(b^2 - a^2)e^{-10b}$$

To calculate the density of $X - Y$, we simply have to calculate the density of the first component a of $g(X, Y)$ which can be done by calculating the marginal density of $g(X, Y)$:

$$f_{X-Y}(a) = \int_{-\infty}^{\infty} f_{g(X,Y)}(a, b)db = \int_{b=|a|}^{\infty} \frac{625}{4}(b^2 - a^2)e^{-10b}db$$

as $f_{g(X,Y)}(a, b)$ is 0 outside of E. Thus, as calculated by Wolfram Alpha:

$$f_{X-Y}(a) = \frac{5}{2}(10|a| + 1)e^{-10|a|}$$

4.0.1 Transformed Variables

Recall that a multivariate continuous random variable is a mapping from the sample space $\mathcal{S}$ to the set of vectors in $\mathbb{R}^n$: $X : \mathcal{S} \to \mathbb{R}^n$; let us call D the set of possible values of $X(s)$. As an example, D could be the set of (x, y) with $y > 0$.

Suppose that we have a mapping $g : D \to E$ where E is a subet of $\mathbb{R}^n$: we call the transformed random variable $g \circ X$, often written as $g(X)$, the random variable whose values are $g(X(s))$ for each outcome s in the sample space..

This is the same as saying: suppose that we have a series of continuous random variables X_i $(0 \leq i \leq n)$ and a series of functions $g_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ we consider the transformed variable $g(X)$ defined by:

$$g(X)(s) = (g_1(X_1(s), \ldots, X_n(s)), \ldots, g_n(X_1(s), \ldots, X_n(s))).$$

Suppose that g is a derivable transformation (which means: $g_i(x_1, ..., x_n)$ is derivable, i.e. $\frac{\partial g_i}{\partial x_j}$ exists for each $0 \leq i, j \leq n$) and that the inverse $g^{-1} : E \longrightarrow D$ of g exists. The transformation theorem, proved in [HMC20, 2.7], says that:

- $g(X)$ It is a continuous random variable.
- if X has a probability density function (pdf) $f_X : D \to \mathbb{R}$ then the pdf of $g(X)$, noted $f_{g(X)}$, is equal to the following, $\forall \mathbf{e} \in E$:

$$f_{g(x)}(\mathbf{e}) = f_X(g^{-1}(\mathbf{e})) \cdot |J_{g^{-1}}(\mathbf{e})|$$

Where $J_{g^{-1}}(\mathbf{e})$ is the Jacobian of the transformation g^{-1} and is non-zero at least in on $\mathbf{e}$: The Jacobian is the determinant of the matrix of each partial derivative of g^{-1}:

$$J_{g^{-1}}(\mathbf{e}) = \begin{vmatrix} \frac{\partial g_1}{\partial e_1}(\mathbf{e}) & \frac{\partial g_1}{\partial e_2}(\mathbf{e}) & \cdots & \frac{\partial g_1}{\partial e_n}(\mathbf{e}) \\ \frac{\partial g_2}{\partial e_1}(\mathbf{e}) & \frac{\partial g_2}{\partial e_2}(\mathbf{e}) & \cdots & \frac{\partial g_2}{\partial e_n}(\mathbf{e}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial e_1}(\mathbf{e}) & \frac{\partial g_n}{\partial e_2}(\mathbf{e}) & \cdots & \frac{\partial g_n}{\partial e_n}(\mathbf{e}) \end{vmatrix}$$

It is important to note that D (the value-set of X and source set of the transformation) and E (the value-set of the transformation) are rarely equal to complete $\mathbb{R}^n$. E.g. It could be a subset of a plane where the lowest x depends on y as we shall see in the exercise.

Coming back to our example random variable $T(s) = X(s) - Y(s)$ with $X \sim \Gamma(2,5)$ and $Y \sim \Gamma(2,5)$ independent. We consider the random variable couple $C(s) = (X(s), Y(s))$ which gives a point for each possible random outcome (the points are in the upper-half quadrant $\{(x,y)|x > 0 \text{ and } y > 0\}$).

To calculate the density of the difference we use the following transformation:

$$\begin{array}{rlcl} g: & \mathbb{R}^+ \times \mathbb{R}^+ & \longrightarrow & E \\ & (x,y) & \mapsto & (a,b) = g(x,y) = (x-y, x+y) \end{array}$$

What is the destination subset $E \subset \mathbb{R} \times \mathbb{R}$ of the transformation g? It is the set of (a,b) such that we can find (x,y) with $x - y = a$ and $x + y = b$. This can be done by solving the equation to obtain x and y from a and b:

$$\left\{ \begin{array}{l} a = x - y \\ b = x + y \end{array} \right. \iff \left\{ \begin{array}{l} a = x - y \\ a + b = 2x \end{array} \right.$$

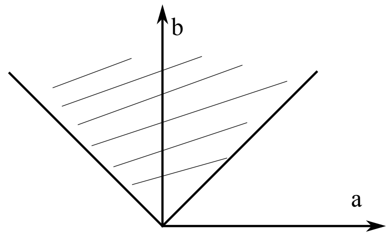

Figure 2: The set of (a, b) reached by the transformation g.

$$\iff \begin{cases} y = \frac{1}{2}(a+b) - a \\ x = \frac{1}{2}(a+b) \end{cases} \iff \begin{cases} y = \frac{1}{2}(b-a) \\ x = \frac{1}{2}(a+b) \end{cases}$$

Thus, E is the set of pairs $(a, b) \in \mathbb{R} \times \mathbb{R}$ such that $b - a > 0$ and $b + a > 0$ which his the hatched region in Figure 2. From solving this equation, we can right away calculate the inverse transformation g^{-1}:

$$(x, y) = g^{-1}(a, b) = \left(\frac{1}{2}(a+b), \frac{1}{2}(b-a)\right)$$

We now apply the transformation theorem above with the transformation g so as to get the density of $g(C)$. This will give us the joint density of $(X-Y, X+Y)$ from which we shall be able to deduce the density of $X - Y$ as intended.

To apply the theorem, we need to calculate the Jacobian of g^{-1}:

$$J_{g^{-1}}(a, b) = \begin{vmatrix} \frac{\partial(\frac{1}{2}(b-a))}{\partial a} & \frac{\partial(\frac{1}{2}(b-a))}{\partial b} \\ \frac{\partial(\frac{1}{2}(b+a))}{\partial a} & \frac{\partial(\frac{1}{2}(b+a))}{\partial b} \end{vmatrix} = \frac{1}{2} \cdot \begin{vmatrix} -1 & 1 \\ 1 & 1 \end{vmatrix} = \frac{1}{2} \cdot (-2) = -1$$

Let us remind that the density of the X and Y (both $\sim \Gamma(2, 5)$) is given by the function $f(t) = \frac{25}{1} \cdot t^{2-1} \cdot e^{-5t} = 25te^{-5t}$ and they are independent thus the joint density of (X, Y) is:

$$f_{(X,Y)}(x, y) = 25xe^{-5x} \cdot 25ye^{-5y} = 625 \cdot x \cdot y \cdot e^{-5(x+y)}$$

The theorem thus gives us the joint density of $g(X, Y)$ is the following for $(a, b) \in E$ and is 0 otherwise:

$$f_{g(X,Y)}(a, b) = f_{(X,Y)}(g^{-1}(a, b)) \cdot |J_{g^{-1}}(a, b)| = \frac{625}{4} \cdot (b-a) \cdot (a+b) \cdot e^{-5 \cdot 2 \cdot b} = \frac{625}{4}(b^2 - a^2)e^{-10b}$$

To calculate the density of $X - Y$, we simply have to calculate the density of the first component a of $g(X, Y)$ which can be done by calculating the marginal density of $g(X, Y)$:

$$f_{X-Y}(a) = \int_{-\infty}^{\infty} f_{g(X,Y)}(a,b)db = \int_{b=|a|}^{\infty} \frac{625}{4}(b^2 - a^2)e^{-10b}db$$

as $f_{g(X,Y)}(a,b)$ is 0 outside of E. Thus, as calculated by Wolfram Alpha:

$$f_{X-Y}(a) = \frac{5}{2}(10|a| + 1)e^{-10|a|}$$

References

[Ano21] Anonymous. Gamma distribution. web-page, February 2021. See `https://en.wikipedia.org/wiki/Gamma_distribution`.

[HMC20] Robert Hogg, Joseph McKean, and Allen Craig. *Introduction to Mathematical Statistics*. Pearson Education, eigth edition (global) edition, 2020.

[inc20] Wolfram Alpha inc. Wolfram alpha, 2020. See `https://wolframalpha.com`.

[LCH21] Maxwell W. Libbrecht, Rachel C. W. Chan, and Michael M. Hoffman. Segmentation and genome annotation algorithms. *CoRR*, abs/2101.00688, 2021.