8 Special distributions

8a Standard normal distributions: 1-dim and 2-dim

The most important one-dimensional distribution is, of course, the standard normal distribution, $N(0,1)$:

(8a1) $$X \sim N(0,1) \iff f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} .$$

It is not evident that the normalizing constant is indeed $1/\sqrt{2\pi}$. We'll see soon that it becomes evident when considering 2-dim normal distribution! Moreover, many properties of the 1-dim normal distribution are rooted in higher dimensions.

The standard two-dimensional normal distribution $N(0,1) \otimes N(0,1)$,

(8a2) $$(X,Y) \sim N(0,1) \otimes N(0,1) \iff f_{X,Y}(x,y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2} ,$$

describes a pair of independent standard normal random variables. The density at a point (x,y) depends only on its distance $r = \sqrt{x^2+y^2}$ from the origin, which is a very special property of the normal distribution. Turn to polar coordinates:

$$X = R\cos\Phi , \qquad Y = R\sin\Phi ,$$
$$f_{R,\Phi}(r,\varphi) = r f_{X,Y}(r\cos\varphi, r\sin\varphi) = r\frac{1}{2\pi} e^{-r^2/2} = f_R(r) f_\Phi(\varphi) ,$$
$$\text{where} \quad f_R(r) = re^{-r^2/2} \quad \text{and} \quad f_\Phi(\varphi) = \frac{1}{2\pi} .$$

We see that R and Φ are independent, and $\Phi \sim U(0,2\pi)$. Note that $\int_0^\infty re^{-r^2/2}\,dr = \int_0^\infty e^{-r^2/2}\,d(r^2/2) = \int_0^\infty e^{-u}\,du = 1$ and, of course, $\int_0^{2\pi} \frac{1}{2\pi}\,d\varphi = 1$. This is the promized explanation of the constant $\frac{1}{2\pi}$ in (8a2) as well as $1/\sqrt{2\pi}$ in (8a1).

When doing computerized simulations, one uses a random generator that produces random (or rather, pseudo-random) numbers distributed uniformly on $(0,1)$; other distributions must be produced out of that. In order to transform $U(0,1)$ into $N(0,1)$ one may use the quantile function of the normal distribution, which is difficult, since it is the inverse function to the non-elementary function $x \mapsto \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2}\,du$. Strangely enough, it is easier to transform $U(0,1) \otimes U(0,1)$ into $N(0,1) \otimes N(0,1)$ than $U(0,1)$ into $N(0,1)$. Just take

(8a3) $$\begin{aligned} R &= \sqrt{2\ln\frac{1}{U}} , \qquad & X &= R\cos\Phi , \\ \Phi &= 2\pi V , & Y &= R\sin\Phi , \end{aligned}$$

and you get

$$(U,V) \sim U(0,1) \otimes U(0,1) \implies (X,Y) \sim N(0,1) \otimes N(0,1) .$$

Independence of R and Φ implies independence of, say, R^2 and $\cos^2\Phi$, that is,

(8a4) $$(X,Y) \sim N(0,1) \otimes N(0,1) \implies X^2+Y^2 \text{ and } \frac{X^2}{X^2+Y^2} \text{ are independent.}$$

Of course, you can check it by calculating the transformed density via the Jacobian of the transformation $T(x,y) = \big(x^2+y^2, x^2/(x^2+y^2)\big)$. However, could you guess, which functions of X, Y are independent?

8a5 Exercise. Find the distribution of $X^2/(X^2+Y^2)$, and also of X^2/Y^2, for $(X,Y) \sim N(0,1) \otimes N(0,1)$.

Hint. These are $\cos^2 \Phi$ and $1/\tan^2 \Phi$ where $\Phi \sim U(0, 2\pi)$.

8b Gamma distribution

First, recall the standard exponential distribution:

$$X \sim \operatorname{Exp}(1) \quad \Longleftrightarrow \quad F_X(x) = 1 - e^{-x} \text{ for } x \in [0,\infty) \quad \Longleftrightarrow \quad f_X(x) = e^{-x} \cdot \mathbf{1}_{(0,\infty)}(x)\,.$$

Now, return to $(X,Y) \sim N(0,1) \otimes N(0,1)$ and calculate the distribution of $X^2 + Y^2 = R^2$:

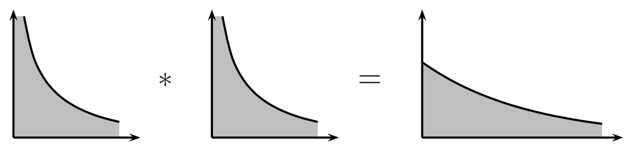
$$F_{R^2}(u) = \mathbb{P}\left(R^2 \le u \right) = \mathbb{P}\left(R \le \sqrt{u} \right) = \int_0^{\sqrt{u}} r e^{-r^2/2}\,dr = -e^{-r^2/2}\Big|_{r=0}^{r=\sqrt{u}} = 1 - e^{-u/2}\,.$$

So,

$$(X,Y) \sim N(0,1) \otimes N(0,1) \quad \Longrightarrow \quad \frac12 (X^2+Y^2) \sim \operatorname{Exp}(1)\,. \tag{8b1}$$

We see that an exponential random variable is a sum of two independent, identically distributed random variables, $\frac12 X^2$ and $\frac12 Y^2$. It is intriguing! What is their distribution? The cumulative distribution function is non-elementary, but the densiity is elementary:

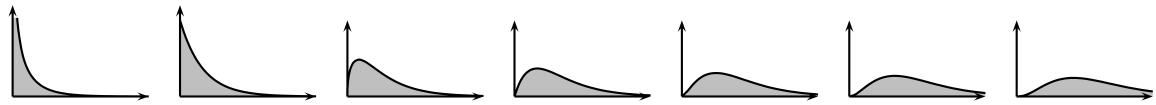
$$\begin{aligned}
&X \sim N(0,1)\,, \quad U = \frac12 X^2\,;\\
&u = \frac12 x^2 \quad \Longrightarrow \quad \frac{du}{dx} f_U(u) = 2 f_X(x)\,; \qquad \text{(why the “2”?)}\\
&f_U(u) = \frac{2}{\sqrt{2u}} f_X(\sqrt{2u}) = \frac{2}{\sqrt{2u}} \frac{1}{\sqrt{2\pi}} e^{-(\sqrt{2u})^2/2} = \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{u}} e^{-u}\,.
\end{aligned}$$



The latter can be written as $\operatorname{Gamma}(1/2) * \operatorname{Gamma}(1/2) = \operatorname{Gamma}(1)$ after defining *gamma distributions* as follows.[1]

$$X \sim \operatorname{Gamma}(\alpha) \quad \Longleftrightarrow \quad f_X(x) = \operatorname{const}(\alpha) \cdot x^{\alpha-1} e^{-x} \cdot \mathbf{1}_{(0,\infty)}(x)\,. \tag{8b2}$$

[1] Another notation $\Gamma(\alpha,\lambda)$ is in use; $X \sim \Gamma(\alpha,\lambda) \iff \lambda X \sim \Gamma(\alpha,1) \iff \lambda X \sim \operatorname{Gamma}(\alpha)$; the second parameter λ is just scaling. For integer n, the distribution $\Gamma\big(\frac n2, \frac12\big)$ is known also as the distribution χ^2_n.



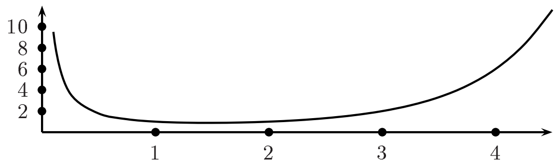
$\alpha = 0.5$ $\alpha = 1$ $\alpha = 1.5$ $\alpha = 2$ $\alpha = 2.5$ $\alpha = 3$ $\alpha = 3.5$

(We'll see soon, why $\alpha - 1$ is more convenient than α.) The function is integrable if and only if $\alpha > 0$; the latter will be assumed. The normalization constant $\operatorname{const}(\alpha)$ is $1/\Gamma(\alpha)$, where Γ is so-called gamma-function:

$$\begin{aligned} &\operatorname{const}(\alpha) = \frac{1}{\Gamma(\alpha)}\,; \\ &\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\,dx\,; \\ &\Gamma(1) = 1\,; \qquad \Gamma(1/2) = \sqrt{\pi}\,; \qquad\qquad \Big(\text{recall } \frac{1}{\sqrt{\pi}}\frac{1}{\sqrt{u}}e^{-u}\Big) \\ &\Gamma(\alpha+1) = \alpha\Gamma(\alpha)\,. \end{aligned} \tag{8b3}$$

The latter is achieved by integrating by parts: $\Gamma(\alpha+1) = \int_0^\infty x^\alpha e^{-x}\,dx = -x^\alpha e^{-x}\big|_0^\infty + \int_0^\infty \alpha x^{\alpha-1} e^{-x}\,dx = \alpha\Gamma(\alpha)$. It follows that $\Gamma(2) = 1\cdot\Gamma(1) = 1$; $\Gamma(3) = 2\Gamma(2) = 2$; $\Gamma(4) = 3\Gamma(3) = 6$; and so on. Similarly, $\Gamma(1.5) = \frac12\Gamma\big(\frac12\big) = \frac12\sqrt{\pi}$; $\Gamma(2.5) = 1.5\Gamma(1.5) = \frac34\sqrt{\pi}$; and so on:

$$\begin{aligned} &\Gamma(n+1) = n! \\ &\Gamma\big(n+\tfrac12\big) = \tfrac{1\cdot3\cdot5\cdot\dots\cdot(2n-1)}{2^n}\sqrt{\pi}\,. \end{aligned} \tag{8b4}$$



When α is an integer, the cumulative distribution function of $\operatorname{Gamma}(\alpha)$ is elementary and, unexpectedly, is closely related to the Poisson distribution.[2]

In the following subsection we'll see that $\operatorname{Gamma}(\alpha) * \operatorname{Gamma}(\beta) = \operatorname{Gamma}(\alpha+\beta)$.

8c Convolution of gamma distributions, and beta distributions

We may reformulate (8a4) in terms of random variables $U = \frac12 X^2$, $V = \frac12 Y^2$ distributed gamma:

$$(U,V) \sim \operatorname{Gamma}(\tfrac12) \otimes \operatorname{Gamma}(\tfrac12) \quad\Longrightarrow\quad U+V \text{ and } \frac{U}{U+V} \text{ are independent.} \tag{8c1}$$

---

[2]Namely, for $\alpha = n+1$

$$\frac{1}{n!}\int_0^\lambda x^n e^{-x}\,dx = 1 - \sum_{k=0}^n \frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=n+1}^\infty \frac{\lambda^k}{k!}e^{-\lambda}\,,$$

since

$$\begin{aligned} \frac{d}{d\lambda}\bigg(e^{-\lambda}\Big(1+\lambda+\dots+\frac{\lambda^n}{n!}\Big)\bigg) &= e^{-\lambda}\frac{d}{d\lambda}\Big(1+\lambda+\dots+\frac{\lambda^n}{n!}\Big) + \Big(\frac{d}{d\lambda}e^{-\lambda}\Big)\Big(1+\lambda+\dots+\frac{\lambda^n}{n!}\Big) = \\ &= e^{-\lambda}\Big(1+\dots+\frac{\lambda^{n-1}}{(n-1)!}\Big) - e^{-\lambda}\Big(1+\lambda+\dots+\frac{\lambda^n}{n!}\Big) = -\frac{\lambda^n}{n!}e^{-\lambda}\,. \end{aligned}$$

Recall that (8a4) emerged from a remarkable property of $N(0,1)\otimes N(0,1)$: its density is constant on each circle $x^2+y^2=\text{const}$. Similarly, (8c1) emerges from a remarkable property of $\text{Gamma}(1/2)\otimes\text{Gamma}(1/2)$, and even $\text{Gamma}(\alpha)\otimes\text{Gamma}(\beta)$; the density is a function of sum, multiplied by a function of fraction:

$$
\begin{aligned}
&(X,Y)\sim\text{Gamma}(\alpha)\otimes\text{Gamma}(\beta) \quad\Longrightarrow\\
&f_{X,Y}(x,y)=\text{const}\cdot x^{\alpha-1}e^{-x}\cdot y^{\beta-1}e^{-y}=\text{const}\left(\frac{x}{x+y}\right)^{\alpha-1}\left(\frac{y}{x+y}\right)^{\beta-1}(x+y)^{\alpha+\beta-2}e^{-(x+y)}\,.
\end{aligned}
$$

Guessing that $X+Y$ and $\frac{X}{X+Y}$ are independent, we perform the transformation:

$$
\begin{gathered}
\begin{aligned}
u&=x+y; & \qquad x&=uv;\\
v&=\frac{x}{x+y}; & \qquad y&=u(1-v);
\end{aligned}\\
T(x,y)=(u,v);\qquad T:(0,\infty)\times(0,\infty)\to(0,\infty)\times(0,1)\quad\text{is smooth;}\\
\frac{\partial(x,y)}{\partial(u,v)}=\begin{vmatrix} v & u\\ 1-v & -u\end{vmatrix}=-uv-u(1-v)=-u\,;\\
\begin{aligned}
T(x,y)=(u,v)\quad\Longrightarrow\quad f_{U,V}(u,v)&=uf_{X,Y}(x,y)=u\cdot\text{const}\cdot v^{\alpha-1}(1-v)^{\beta-1}\cdot u^{\alpha+\beta-2}e^{-u}=\\
&=\underbrace{\text{const}_1\cdot u^{\alpha+\beta-1}e^{-u}}_{\text{density of Gamma}(\alpha+\beta)}\cdot\underbrace{\text{const}_2\cdot v^{\alpha-1}(1-v)^{\beta-1}}_{\text{density of Beta}(\alpha,\beta)};
\end{aligned}
\end{gathered}
$$

indeed, U and V are independent, and $U\sim\text{Gamma}(\alpha+\beta)$, while the distribution of V is new for us; call it *beta distribution:*

(8c2) $$X\sim\text{Beta}(\alpha,\beta)\quad\Longleftrightarrow\quad f_X(x)=\text{const}\cdot x^{\alpha-1}(1-x)^{\beta-1}\cdot\mathbf{1}_{(0,\infty)}(x)\,.$$

So,
(8c3)
$$(X,Y)\sim\text{Gamma}(\alpha)\otimes\text{Gamma}(\beta)\quad\Longrightarrow\quad\left(X+Y,\frac{X}{X+Y}\right)\sim\text{Gamma}(\alpha+\beta)\otimes\text{Beta}(\alpha,\beta)\,.$$

That is instructive! The simple but non-evident formula

(8c4) $$\text{Gamma}(\alpha)*\text{Gamma}(\beta)=\text{Gamma}(\alpha+\beta)$$

follows immediately from a more complicated but more evident formula (8c3). Once again, many properties of 1-dim distributions are rooted in higher dimensions. Note also that

(8c5) $$(X,Y)\sim\text{Gamma}(\alpha)\otimes\text{Gamma}(\beta)\quad\Longrightarrow\quad\frac{X}{X+Y}\sim\text{Beta}(\alpha,\beta)\,.$$

Beta distribution appears also from order statistics $X^{(1)},\dots,X^{(n)}$ of independent $U(0,1)$ random variables $X_1,\dots,X_n$. Namely,

$$X^{(k)}\sim\text{Beta}(k,n-k+1)\,.$$

Independence stated by (8c3) gives us some conditional distributions:

(8c6) $$(X,Y) \sim \mathrm{Gamma}(\alpha) \otimes \mathrm{Gamma}(\beta) \quad \Longrightarrow \quad \begin{aligned} &\left(X+Y \,\middle|\, \frac{X}{Y} = r \right) \sim \mathrm{Gamma}(\alpha+\beta)\,, \\ &\left(\frac{1}{s} X \,\middle|\, X+Y = s \right) \sim \mathrm{Beta}(\alpha,\beta)\,. \end{aligned}$$

The ratio $U = \frac{X}{Y}$ can be obtained from $V = \frac{X}{X+Y}$ by a simple one-dimensional transformation $U = \frac{V}{1-V}$, $V = \frac{U}{1+U}$. Knowing beta distribution, we may calculate easily the distribution of X/Y, known in statistics as Snedekor's F-distribution, $F_{m,n}$. More exactly,

(8c7) $$(X,Y) \sim \mathrm{Gamma}(\tfrac{m}{2}) \otimes \mathrm{Gamma}(\tfrac{n}{2}) \quad \Longrightarrow \quad \frac{n}{m}\frac{X}{Y} \sim F_{m,n}\,.$$

In order to find the constant in (8c2) we should calculate $\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx$, which is difficult. Instead we note that the normalization constant $\frac{1}{\Gamma(\alpha)} \cdot \frac{1}{\Gamma(\beta)}$ of the distribution $\mathrm{Gamma}(\alpha) \otimes \mathrm{Gamma}(\beta)$ is equal to the normalization constant $\frac{1}{\Gamma(\alpha+\beta)}$ of the distribution $\mathrm{Gamma}(\alpha+\beta)$ multiplied by the normalization constant of the distribution $\mathrm{Beta}(\alpha,\beta)$. Therefore, the latter is equal to $\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$, and the difficult integral is painlessly calculated:

$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}\,;$$

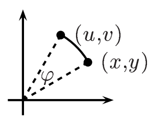
in particular,

$$\int_0^1 x^m (1-x)^n\,dx = \frac{m!n!}{(m+n)!} = \frac{1}{\binom{m+n}{m}}\,.$$

8d Multivariate normal distribution

Let X,Y be independent $N(0,1)$ random variables, then $f_{X,Y}(x,y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2}$ is rotation invariant:

(8d1) $$f_{X,Y}\big(\underbrace{x\cos\varphi - y\sin\varphi}_{u}, \underbrace{x\sin\varphi + y\cos\varphi}_{v}\big) = f_{X,Y}(x,y)$$
$$\text{since } x^2+y^2 = u^2+v^2\,.$$



The rotation is a smooth transformation $\mathbb{R}^2 \to \mathbb{R}^2$ with the Jacobian $J(x,y) = 1$. Thus,

(8d2) $$\begin{cases} U = X\cos\varphi - Y\sin\varphi \\ V = X\sin\varphi + Y\cos\varphi \end{cases} \quad \Longrightarrow \quad f_{U,V} = f_{X,Y}\,,$$

which means that U,V are independent $N(0,1)$ random variables, too!

In particular, $X\cos\varphi - Y\sin\varphi \sim N(0,1)$; in other words,

$$\frac{aX+bY}{\sqrt{a^2+b^2}} \sim N(0,1)\,, \quad \text{that is,} \quad aX+bY \sim N\big(0, a^2+b^2\big)$$

for any $a, b \in \mathbb{R}$ except for $(0,0)$. The second formula holds also for $a = b = 0$ after defining $N(0,0)$ as the degenerate distribution, concentrated at 0.

The set $G_2 = \{aX + bY : a, b \in \mathbb{R}\}$ is a 2-dimensional linear space of random variables. It is an Euclidean space: each $Z \in G_2$ has its norm $\|Z\|$ defined by $\|Z\|^2 = \mathbb{E}(Z^2) = \mathrm{Var}(Z)$, and the norm is generated by the scalar product $(Z_1, Z_2) = \mathbb{E}(Z_1 Z_2) = \mathrm{Cov}(Z_1, Z_2)$. All normalized elements of G_2 are identically distributed:

$$Z \in G_2\,,\ \|Z\| = 1 \quad \Longrightarrow \quad Z \sim N(0,1)\,.$$

In general, two random variables with equal mean values and variances need not be identically distributed. However, within G_2 they need!

Any orthogonal vectors of G_2 are *uncorrelated* random variables: $\mathrm{Cov}(Z_1, Z_2) = (Z_1, Z_2) = 0$. In general, uncorrelated does not mean independent. However, within G_2 it does! Indeed, any pair of orthogonal vectors of G_2 is of the form $\mathrm{const}_1 \cdot U$, $\mathrm{const}_2 \cdot V$ with U, V as in (8d2).

Any orthogonal basis of G_2 may serve equally well as the initial pair (X, Y).

The two-dimensional theory can be generalized for n dimensions.

8d3 Definition. An n-dimensional Gaussian space is a set of random variables (on a given probability space) of the form

$$G_n = \{c_1 X_1 + \dots + c_n X_n : c_1, \dots, c_n \in \mathbb{R}\}$$

where $X_1, \dots, X_n$ are independent $N(0,1)$ random variables.

Clearly, G_n is an n-dimensional Euclidean space, and $\{X_1, \dots, X_n\}$ is its orthonormal basis. Any other basis can be obtained by an n-dimensional rotation (that is, a linear norm preserving transformation). The rotation has the Jacobian $J = \pm 1$, and preserves the n-dimensional density

$$f_{X_1,\dots,X_n}(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2}} e^{-(x_1^2 + \dots + x_n^2)/2}\,.$$

It follows that *any* orthonormal basis of G_n consists of n independent $N(0,1)$ random variables. Therefore, any m-dimensional linear subspace of G_n is an m-dimensional Gaussian space. In particular, $(m = 1)$ any element of G_n is a normally distributed random variable, and $(m = 2)$ any two orthogonal elements of G_n are independent random variables. Within G_n, uncorrelated means independent!

8d4 Definition. (a) Random variables $X_1, \dots, X_n : \Omega \to \mathbb{R}$ have a multivariate normal distribution, if there are $\mu_1, \dots, \mu_n$ such that $X_1 - \mu_1, \dots, X_n - \mu_n$ belong to a Gaussian space.

(b) The rank of the multivariate normal distribution is the dimension of the Gaussian space spanned by $X_1 - \mu_1, \dots, X_n - \mu_n$.

Clearly, $\mu_k = \mathbb{E}(X_k)$.

8d5 Example. Let X, Y, Z have a multivariate normal distribution, and

$$X, Y, Z \sim N(0,1)\,,$$
$$\operatorname{Cov}(X,Y) = \operatorname{Cov}(Y,Z) = \operatorname{Cov}(Z,X) = -\frac{1}{2}\,.$$

What is the rank?

The following orthogonalization process is well-known in linear algebra. We may take the first basis vector $X_1 = X$, since $\|X\| = 1$. The first coefficient of Y is $(Y, X_1) = \operatorname{Cov}(Y, X) = -\frac{1}{2}$, and we may choose the second basis vector X_2 such that $Y = -\frac{1}{2}X_1 + \text{const}\cdot X_2$. That is, X_2 is the normalized $Y + \frac{1}{2}X_1$. We have $\operatorname{Var}(Y + \frac{1}{2}X_1) = \operatorname{Var}(Y) + \frac{1}{4}\operatorname{Var}(X_1) + \operatorname{Cov}(Y, X_1) = 1 + \frac{1}{4}\cdot 1 + (-\frac{1}{2}) = \frac{3}{4}$, thus

$$\begin{cases} X = X_1, \\ Y = -\dfrac{1}{2}X_1 + \dfrac{\sqrt{3}}{2}X_2; \end{cases} \qquad \begin{cases} X_1 = X, \\ X_2 = \dfrac{2}{\sqrt{3}}(Y + \dfrac{1}{2}X). \end{cases}$$

The first two coefficients of Z are:

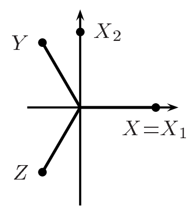
$$\begin{aligned} &(Z, X_1) = \operatorname{Cov}(Z, X) = -\tfrac{1}{2}\,; \\ &(Z, X_2) = \operatorname{Cov}\bigl(Z, \tfrac{2}{\sqrt{3}}(Y + \tfrac{1}{2}X)\bigr) = \tfrac{2}{\sqrt{3}}\operatorname{Cov}(Z,Y) + \tfrac{2}{\sqrt{3}}\cdot\tfrac{1}{2}\operatorname{Cov}(Z,X) = \\ &\quad = \tfrac{2}{\sqrt{3}}\cdot(-\tfrac{1}{2}) + \tfrac{1}{\sqrt{3}}\cdot(-\tfrac{1}{2}) = -\tfrac{\sqrt{3}}{2}\,. \end{aligned}$$

It means that $Z = -\frac{1}{2}X_1 - \frac{\sqrt{3}}{2}X_2 + \text{const}\cdot X_3$, and X_3 is the normalized $Z + \frac{1}{2}X_1 + \frac{\sqrt{3}}{2}X_2$. We have

$$\begin{aligned} &\operatorname{Var}\bigl(Z + \tfrac{1}{2}X_1 + \tfrac{\sqrt{3}}{2}X_2\bigr) = \operatorname{Var}(Z + Y + X) = \\ &\quad = \operatorname{Var}(Z) + \operatorname{Var}(Y) + \operatorname{Var}(X) + 2\operatorname{Cov}(Z,Y) + 2\operatorname{Cov}(Z,X) + 2\operatorname{Cov}(Y,X) = \\ &\quad = 1 + 1 + 1 + 2\cdot(-\tfrac{1}{2}) + 2\cdot(-\tfrac{1}{2}) + 2\cdot(-\tfrac{1}{2}) = 0\,. \end{aligned}$$

It means that our Gaussian space is 2-dimensional, thus, the rank is 2. So,

$$\begin{cases} X = X_1 \\ Y = -\tfrac{1}{2}X_1 + \tfrac{\sqrt{3}}{2}X_2 \\ Z = -\tfrac{1}{2}X_1 - \tfrac{\sqrt{3}}{2}X_2 \end{cases} \qquad \begin{cases} X_1, X_2 \sim N(0,1) \\ X_1, X_2 \text{ are independent} \end{cases}$$



The orthogonalization process pertains to linear algebra, but the following conclusion pertains to probability theory: the joint distribution $P_{X,Y,Z}$ is uniquely determined by the given mean values, variations and covariations. If X', Y', Z' satisfy the same conditions, then $X' = X'_1$, $Y' = -\frac{1}{2}X'_1 + \frac{\sqrt{3}}{2}X'_2$, $Z' = -\frac{1}{2}X'_1 - \frac{\sqrt{3}}{2}X'_2$ with independent $X'_1, X'_2 \sim N(0,1)$. We have $(X_1, X_2) \sim (X'_1, X'_2)$, therefore $(X, Y, Z) \sim (X', Y', Z')$.

In general, the same argument shows that a multivariate normal distribution is uniquely determined by its mean vector and covariance matrix,

$$\big(\mathbb{E}\,(X_k)\big)_{k=1\dots n}\,, \qquad\qquad \big(\operatorname{Cov}(X_k, X_l)\big)_{k,l=1\dots n}\,.$$

If the rank r is equal to the dimension n, then the multivariate normal disribution is absolutely continuous (has an n-dimensional density). If $r = 0$, then the distribution is discrete: a single atom. If $0 < r < n$, then the distribution is singular (concentrated on an r-dimensional subspace of $\mathbb{R}^n$).

Rotation invariance is very useful when dealing with normal distributions.

8d6 Example. Let $(X, Y) \sim N(0, \sigma^2) \otimes N(0, \sigma^2)$, $\sigma > 0$; find the distribution of

$$\frac{X + Y}{|X - Y|}\,.$$

Solution. First, we may assume that $\sigma = 1$, since $(X+Y)/|X-Y| = \big(\frac{1}{\sigma}X + \frac{1}{\sigma}Y\big)/|\frac{1}{\sigma}X - \frac{1}{\sigma}Y|$. Second, rotation by 45° transforms $(X + Y)/|X - Y|$ into $X/|Y|$.[3] Third, use polar coordinates: $X = R\cos\Phi$, $Y = R\sin\Phi$, then $\Phi \sim U(0, 2\pi)$ and $X/|Y| = \cos\Phi/|\sin\Phi|$, which is $1/\tan\Phi$ for $\Phi \in (0, \pi)$ and $-1/\tan\Phi$ for $\Phi \in (\pi, 2\pi)$.

Answer. Cauchy distribution (with the density $\frac{1}{\pi}\frac{1}{1+u^2}$ and the cumulative distribution function $\frac{1}{2} + \frac{1}{\pi}\arctan u$).

In higher dimensions, rotation invariance is even more useful; all you need is some simple geometry.

8d7 Example. Let X, Y, Z be independent random variables distributed $N(0, \sigma^2)$, $\sigma > 0$; find the distribution of

$$\frac{X + Y + Z}{\sqrt{\Big(X - \frac{X+Y+Z}{3}\Big)^2 + \Big(Y - \frac{X+Y+Z}{3}\Big)^2 + \Big(Z - \frac{X+Y+Z}{3}\Big)^2}}\,.$$

Solution. First, we may assume that $\sigma = 1$. Second, in the 3-dim space $\mathbb{R}^3$ the point $\big(\frac{x+y+z}{3}, \frac{x+y+z}{3}, \frac{x+y+z}{3}\big)$ is the orthogonal projection of the point (x, y, z) onto the straight line $\{(u, u, u) : u \in \mathbb{R}\}$; the denominator $\sqrt{\big(X - \frac{X+Y+Z}{3}\big)^2 + \big(Y - \frac{X+Y+Z}{3}\big)^2 + \big(Z - \frac{X+Y+Z}{3}\big)^2}$ is just the distance of the point (x, y, z) from the line; and $x+y+z$ is the signed distance of the projection from the origin, multiplied by $\sqrt{3}$. We rotate $\mathbb{R}^3$, turning the line into, say, the first coordinate axis $\{(u, 0, 0) : u \in \mathbb{R}\}$. Now, the projection of the point (x, y, z) onto the line is $(x, 0, 0)$; the distance of the point (x, y, z) from the line is $\sqrt{y^2 + z^2}$; and the signed distance of the projection from the origin is x. Therefore the given ratio turns into

$$\frac{\sqrt{3}X}{\sqrt{Y^2 + Z^2}}\,.$$

[3] In other words: consider $U = (X + Y)/\sqrt{2}$, $V = (X - Y)/\sqrt{2}$, then $(U, V) \sim N(0, 1) \otimes N(0, 1)$ and $(X + Y)/|X - Y| = U/|V|$ is distributed like $X/|Y|$.

This random variable is symmetric,[4] and its square is, up to a coefficient, of the form (8c7); $\frac{1}{2}X^2 \sim \text{Gamma}(\frac{1}{2})$, $\frac{1}{2}Y^2 + \frac{1}{2}Z^2 \sim \text{Gamma}(1)$, these being independent, therefore

$$\frac{2}{1} \cdot \frac{\frac{1}{2}X^2}{\frac{1}{2}Y^2 + \frac{1}{2}Z^2} \sim F_{1,2} \,.$$

[4]That is, $\frac{\sqrt{3}X}{\sqrt{Y^2+Z^2}}$ and $-\frac{\sqrt{3}X}{\sqrt{Y^2+Z^2}}$ are identically distributed, which follows from the fact that triples (X,Y,Z) and $(-X,Y,Z)$ are identically distributed.