Distributions related to normal distribution

How to get PDF

From CDF to PDF

$$P(X \le x) \quad \overset{\text{Differentiate}}{\Rightarrow} \quad f_X(x)$$

From Jacobian to PDF

$$\begin{aligned} f_Y(y) &= f_X(x)\left|\frac{dx}{dy}\right| \\ f_{U,V}(u,v) &= f_{X,Y}(x,y)\left|\frac{\partial(x,y)}{\partial(u,v)}\right| \\ f_{Y_1,\ldots,Y_n}(y_1,\ldots,y_n) &= f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)\left|\frac{\partial(x_1,\ldots,x_n)}{\partial(y_1,\ldots,y_n)}\right| \end{aligned}$$

where

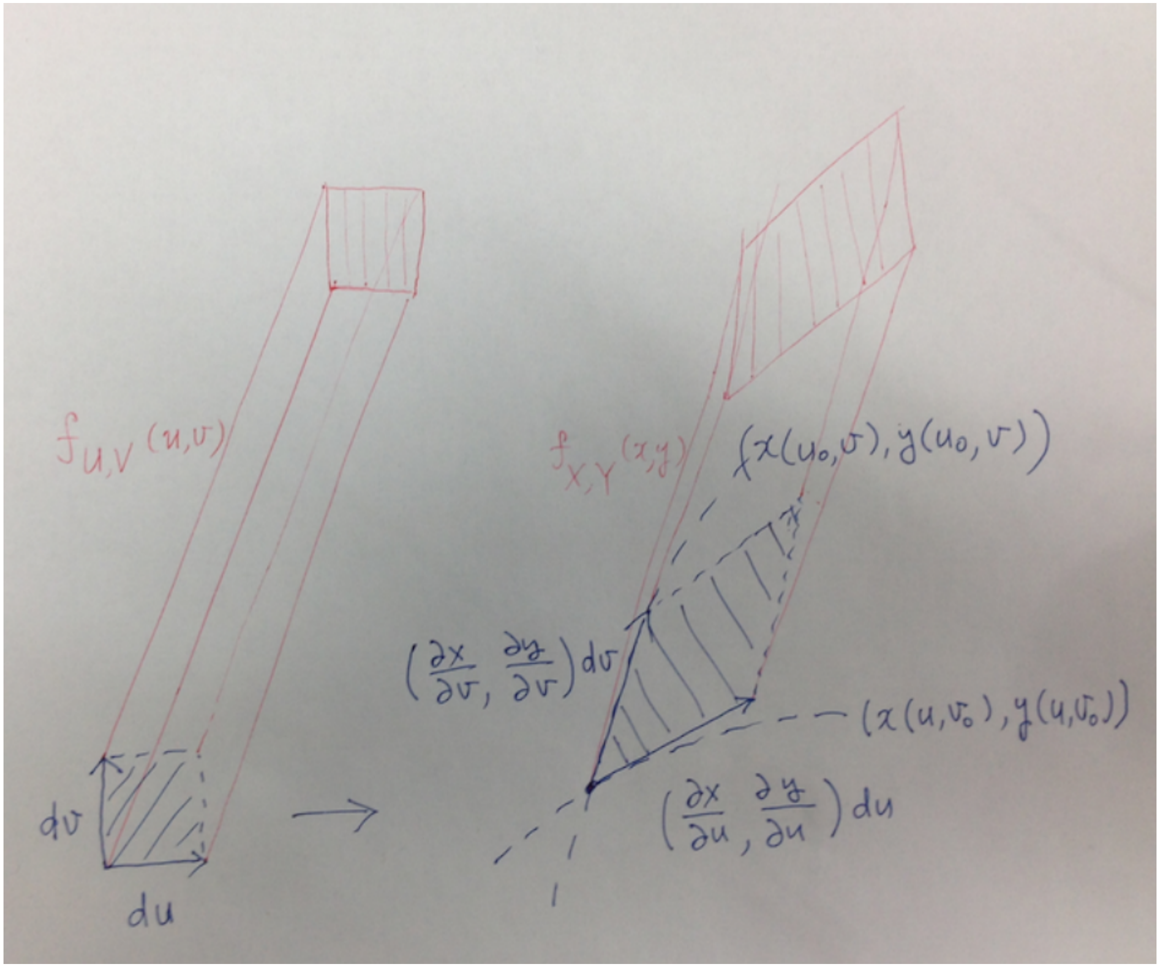
$$\left|\frac{\partial(x_1,\ldots,x_n)}{\partial(y_1,\ldots,y_n)}\right| = \left| det \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{pmatrix} \right|$$

Property of Jacobian

$$\left|\frac{\partial(x_1,\ldots,x_n)}{\partial(y_1,\ldots,y_n)}\right| = \frac{1}{\left|\frac{\partial(y_1,\ldots,y_n)}{\partial(x_1,\ldots,x_n)}\right|}$$

$$P(Y \le y) = P(X \le x) \quad \Rightarrow \quad f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

$$P(Y \le y) = 1 - P(X \le x) \quad \Rightarrow \quad f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$



Hight on left	$f_{U,V}(u_0, v_0)$
Area on left	$dudv$
Volumn on left	$f_{U,V}(u_0, v_0)dudv$
Hight on right	$f_{X,Y}(x_0, x_0)$
Area on right	$\left\lvert\frac{\partial(x,y)}{\partial(u,v)}\right\rvert dudv$
Volumn on right	$f_{X,Y}(x_0, x_0)\left\lvert\frac{\partial(x,y)}{\partial(u,v)}\right\rvert dudv$

$$f_{U,V}(u,v)dudu = f_{X,Y}(x,y)\left\lvert\frac{\partial(x,y)}{\partial(u,v)}\right\rvert dudv \quad \Rightarrow \quad f_{U,V}(u,v) = f_{X,Y}(x,y)\left\lvert\frac{\partial(x,y)}{\partial(u,v)}\right\rvert$$

Example - PDF of $Y = X^3$, where $X \sim U(0,1)$

Use CDF

For $0 < y < 1$,

$$P(Y \leq y) = P(X \leq y^{1/3}) = y^{1/3} \quad \overset{\text{Differentiate}}{\Rightarrow} \quad f_Y(y) = \frac{1}{3}y^{-2/3} \quad \text{for } 0 < y < 1$$

Use Jacobean

With $y = x^3$, for $0 < y < 1$

$$\frac{dy}{dx} = 3x^2 = 3\left(x^3\right)^{2/3} = 3y^{2/3} \quad \Rightarrow \quad \frac{dx}{dy} = 1/\left(\frac{dy}{dx}\right) = \frac{1}{3}y^{-2/3} \quad \Rightarrow \quad \left|\frac{dx}{dy}\right| = \frac{1}{3}y^{-2/3}$$

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right| = \frac{1}{3}y^{-2/3} \quad \text{for } 0 < y < 1$$

Log-normal distribution Log-$\mathcal{N}(\mu, \sigma^2)$

Normal and Log-normal

$$Y \sim \mathcal{N}(\mu, \sigma^2) \quad \Leftrightarrow \quad X = e^Y \sim \text{Log-}\mathcal{N}(\mu, \sigma^2)$$

PDF, Mean, Variance

PDF	$\dfrac{1}{x\sqrt{2\pi\sigma^2}} e^{-\frac{(\log x - \mu)^2}{2\sigma^2}}$
Mean	$e^{\mu + \frac{1}{2}\sigma^2}$
Variance	$\left(e^{\sigma^2} - 1\right) e^{2\mu + \sigma^2}$

$$\begin{aligned} P(X \le x) = P(Y \le \log x) &= \int_{-\infty}^{\log x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(s-\mu)^2}{2\sigma^2}} ds \\ \Rightarrow \quad f_X(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\log x - \mu)^2}{2\sigma^2}} \cdot \frac{1}{x} \quad \text{for } x > 0 \end{aligned}$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \quad \Rightarrow \quad f_X(x) = f_Y(y) \left| \frac{dy}{dx} \right| = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\log x - \mu)^2}{2\sigma^2}} \cdot \frac{1}{x} \quad \text{for } x > 0$$

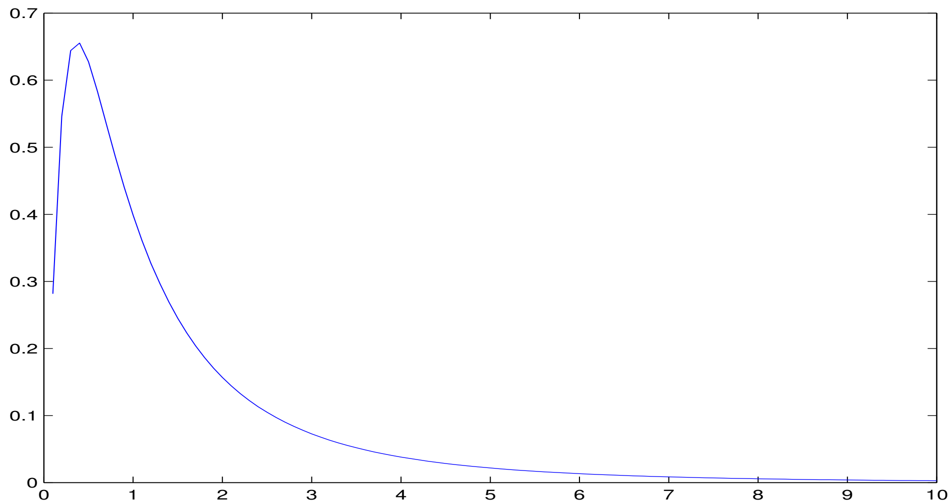


Figure 1: Log-normal distribution.

```
clear all; close all; clc;

x=0.1:0.1:10;
p=pdf('logn',x,0,1);
plot(x,p)
```

Chi-square distribution χ^2_d

Recall - Gamma distribution $\Gamma(\alpha, \lambda)$

$$
\begin{aligned}
&(1) \quad Exp(\lambda) \stackrel{d}{=} \Gamma(1, \lambda) \\
&(2) \quad Exp(\lambda) * Exp(\lambda) \stackrel{d}{=} \Gamma(2, \lambda) \\
&(3) \quad Exp(\lambda) * Exp(\lambda) * \cdots * Exp(\lambda) \stackrel{d}{=} \Gamma(n, \lambda) \\
&(4) \quad \Gamma(\alpha, \lambda) * \Gamma(\beta, \lambda) \stackrel{d}{=} \Gamma(\alpha + \beta, \lambda)
\end{aligned}
$$

Definition - Chi-square distribution χ^2_d

$$\sum_{i=1}^{d} Z_i^2 \sim \chi^2_d \qquad \text{where} \quad Z_i \quad \text{IID} \quad N(0, 1^2)$$

Properties - Chi-square distribution χ^2_d

$$
\begin{aligned}
&(1) \quad \chi^2_1 \stackrel{d}{=} Z_1^2 \stackrel{d}{=} \Gamma\left(\frac{1}{2}, \frac{1}{2}\right) \\
&(2) \quad \chi^2_d \stackrel{d}{=} Z_1^2 + \cdots + Z_d^2 \stackrel{d}{=} \Gamma\left(\frac{1}{2}, \frac{1}{2}\right) * \cdots * \Gamma\left(\frac{1}{2}, \frac{1}{2}\right) \stackrel{d}{=} \Gamma\left(\frac{d}{2}, \frac{1}{2}\right) \\
&(3) \quad \phi_{\chi^2_d}(t) = \left(\frac{1}{\sqrt{1-2t}}\right)^d
\end{aligned}
$$

For $x > 0$

$$P(Z_1^2 \le x) = P(-\sqrt{x} \le Z_1 \le \sqrt{x}) = \int_{-\sqrt{x}}^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds = 2\int_0^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds$$

$$\Rightarrow \quad f_{Z_1^2}(x) = 2\frac{1}{\sqrt{2\pi}} e^{-\frac{x}{2}} \frac{1}{2} x^{-1/2} = \frac{\frac{1}{2}\left(\frac{1}{2}x\right)^{\frac{1}{2}-1} e^{-\frac{1}{2}x}}{\Gamma(\frac{1}{2})} = f_{\Gamma(\frac{1}{2},\frac{1}{2})}(x)$$

$$\Rightarrow \quad (1) \quad \chi_1^2 \stackrel{d}{=} Z_1^2 \stackrel{d}{=} \Gamma\left(\frac{1}{2}, \frac{1}{2}\right)$$

By the property (4) of Gamma distribution

$$(2) \qquad \chi_d^2 \stackrel{d}{=} Z_1^2 + \cdots + Z_d^2 \stackrel{d}{=} \Gamma\left(\frac{1}{2}, \frac{1}{2}\right) * \cdots * \Gamma\left(\frac{1}{2}, \frac{1}{2}\right) \stackrel{d}{=} \Gamma\left(\frac{d}{2}, \frac{1}{2}\right)$$

With $\lambda = \frac{1}{2} - t$

$$\begin{aligned}
(3) \quad \phi_{\chi_d^2}(t) &= \int_0^\infty e^{tx} \frac{\frac{1}{2}\left(\frac{1}{2}x\right)^{\frac{d}{2}-1} e^{-\frac{1}{2}x}}{\Gamma(\frac{d}{2})} dx \\
&= \int_0^\infty \frac{\frac{1}{2}\left(\frac{1}{2}x\right)^{\frac{d}{2}-1} e^{-(\frac{1}{2}-t)x}}{\Gamma(\frac{d}{2})} dx \\
&= \left(\frac{1}{\sqrt{1-2t}}\right)^d \int_0^\infty \underbrace{\frac{\lambda\,(\lambda x)^{\frac{d}{2}-1} e^{-\lambda x}}{\Gamma(\frac{d}{2})}}_{\text{PDF of } \Gamma(\frac{d}{2},\lambda)} dx = \left(\frac{1}{\sqrt{1-2t}}\right)^d
\end{aligned}$$

Mean and variance of geometric, exponential, gamma, to chi-square distribution

	mean	variance
$Geo(p)$	$\frac{1}{p}$	$\frac{q}{p^2}$
$\frac{1}{n}Geo(p)$	$\frac{1}{np}$	$\frac{q}{(np)^2}$
$Exp(\lambda) = \Gamma(1, \lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$\Gamma(n, \lambda)$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$
$\Gamma(\alpha, \lambda)$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$
$\chi^2_1 = \Gamma(\frac{1}{2}, \frac{1}{2})$	$\frac{\frac{1}{2}}{\frac{1}{2}} = 1$	$\frac{\frac{1}{2}}{(\frac{1}{2})^2} = 2$
$\chi^2_d = \Gamma(\frac{d}{2}, \frac{1}{2})$	d	$2d$

Student t distribution t_d

Definition

$$\frac{Z}{\sqrt{\frac{V}{d}}} \sim t_d \qquad \text{where} \qquad Z \sim N(0,1^2) \text{ and } V \sim \chi^2_d \text{ are independent}$$

Why chi-square and student t

For n iid samples X_i from $N(\mu, \sigma^2)$, let $\bar{X}$ and S^2 be the sample mean and variance:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \qquad \text{and} \qquad S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$$

Then,

(1) $\bar{X}$ and S^2 are independent

(2) $$\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \overset{d}{=} N(0,1^2) \qquad \text{and} \qquad \frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma}\right)^2 \overset{d}{=} \chi^2_{n-1}$$

(3) $$\frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} = \frac{\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{\frac{(n-1)S^2}{\sigma^2}}{n-1}}} \overset{(2)}{=} \frac{N(0,1)}{\sqrt{\frac{\chi^2_{n-1}}{n-1}}} \overset{(1)}{=} t_{n-1}$$

PDF

$$f(x) \propto \left(1 + \frac{1}{d}x^2\right)^{-\frac{d+1}{2}} \quad \Rightarrow \quad f(x) = \frac{1}{\sqrt{d}B\left(\frac{1}{2}, \frac{d}{2}\right)} \left(1 + \frac{1}{d}x^2\right)^{-\frac{d+1}{2}}$$

Mean and variance

Mean	0	for $d > 1$
Variance	$\frac{d}{d-2}$	for $d > 2$

Related distribution - Cauchy distribution - $d = 1$

$$f(x) \propto \frac{1}{1+x^2} \quad \Rightarrow \quad f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}$$

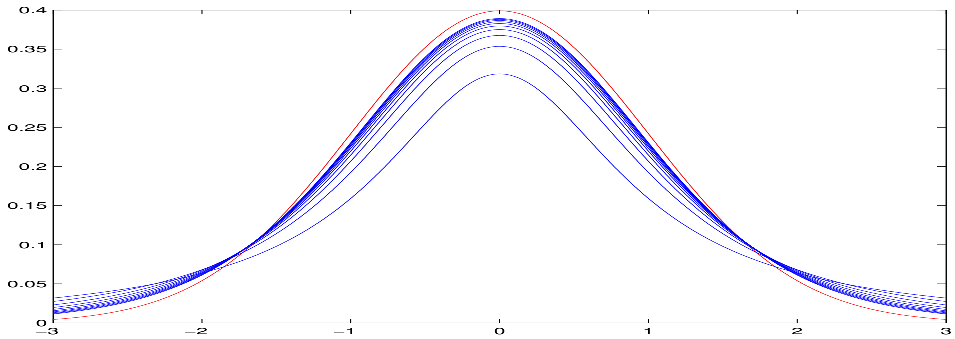


Figure 2: t_d has a fat tail. As $d \to \infty$, t_d converges to $N(0, 1^2)$.

```
clear all; close all; clc;

x=-3:0.01:3;
y=pdf('Normal',x,0,1);
plot(x,y,'-r'); hold on

n=10;
for i=1:n
    y=pdf('T',x,i);
    plot(x,y,'-b');
    pause(0.5)
end
```

Key fact and its consequence

Recall - Multivariate normal $\mathbf{x} \sim \mathcal{N}(\mu, \Sigma)$

(1) μ and Σ completely determine the multivariate normal distribution

(2) If off diagonals of Σ are all 0, then all the components of $\mathbf{x}$ are independent

(3) If for fixed i, $\Sigma_{ij} = 0$ for all $j \neq i$, then $\mathbf{x}_i$ is independent to $\mathbf{x}_j$, $j \neq i$

Key fact

Let X_i be iid with mean μ and variance σ^2. With $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ we have

$$\begin{aligned} Cov(\bar{X}, X_i - \bar{X}) &= Cov(\bar{X}, X_i) - Cov(\bar{X}, \bar{X}) \\ &= Cov\left(\frac{\sum_{j=1}^n X_j}{n}, X_i\right) - Cov\left(\frac{\sum_{j=1}^n X_j}{n}, \frac{\sum_{k=1}^n X_k}{n}\right) \\ &= \frac{1}{n} \cdot \sigma^2 - \frac{1}{n^2} \cdot n\sigma^2 = 0 \end{aligned}$$

Consequence

$$\begin{aligned} & \bar{X}, X_1 - \bar{X}, X_2 - \bar{X}, \cdots, X_n - \bar{X} \text{ are multivariate normal} \\ \Rightarrow\ & \text{Since } Cov(\bar{X}, X_i - \bar{X}) = 0,\ \bar{X} \text{ and } X_1 - \bar{X}, \cdots, X_n - \bar{X} \text{ are independent} \\ \Rightarrow\ & \bar{X} \text{ and } S^2 \text{ are independent} \end{aligned}$$

$$\begin{aligned}
\sum_{i=1}^{n}(X_i-\mu)^2 &= \sum_{i=1}^{n}((X_i-\bar{X})+(\bar{X}-\mu))^2 \\
&= \sum_{i=1}^{n}(X_i-\bar{X})^2+n(\bar{X}-\mu)^2+2(\bar{X}-\mu)\sum_{i=1}^{n}(X_i-\bar{X}) \\
&= \sum_{i=1}^{n}(X_i-\bar{X})^2+n(\bar{X}-\mu)^2
\end{aligned}$$

$$\underbrace{\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2}_{\chi^2_n} = \sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 + \underbrace{\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)^2}_{\chi^2_1}$$

Consequence of key fact

$$\Rightarrow \quad \left(\frac{1}{\sqrt{1-2t}}\right)^n = \phi_{\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2}(t) \cdot \left(\frac{1}{\sqrt{1-2t}}\right)$$

$$\Rightarrow \quad \phi_{\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2}(t) = \left(\frac{1}{\sqrt{1-2t}}\right)^{n-1}$$

$$\Rightarrow \quad \sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 \sim \chi^2_{n-1}$$

With $t = \frac{z}{\sqrt{\frac{v}{r}}}$ and $u = v$, where $Z \sim N(0, 1^2)$ and $V \sim \chi^2_d$ are independent,

$$\left|\frac{\partial(z,v)}{\partial(t,u)}\right| = \left|\frac{\partial(t,u)}{\partial(z,v)}\right|^{-1} = \left|det \begin{pmatrix} \frac{1}{\sqrt{\frac{v}{r}}} & * \\ 0 & 1 \end{pmatrix}\right|^{-1} = \sqrt{\frac{v}{r}}$$

With $\lambda = \frac{1+\frac{t^2}{d}}{2}$,

$$\begin{aligned} f_{T,U}(t,u) &= f_{Z,V}(z,v)\left|\frac{\partial(z,v)}{\partial(t,u)}\right| \\ &= \frac{\frac{1}{2}(\frac{1}{2}v)^{\frac{d}{2}-1}}{\sqrt{2\pi}\Gamma(\frac{r}{2})} e^{-\frac{z^2}{2}} e^{-\frac{1}{2}v} \sqrt{\frac{v}{d}} \\ &= \frac{\frac{1}{2}(\frac{1}{2}u)^{\frac{d}{2}-1}}{\sqrt{2\pi}\Gamma(\frac{d}{2})} e^{-\frac{1+\frac{t^2}{d}}{2}u} \sqrt{\frac{u}{d}} \\ &= \frac{1}{\sqrt{d}B\left(\frac{1}{2},\frac{d}{2}\right)}\left(1+\frac{t^2}{d}\right)^{-\frac{d+1}{2}} \cdot \underbrace{\left[\frac{\lambda\,(\lambda u)^{\frac{d+1}{2}-1} e^{-\lambda u}}{\Gamma(\frac{r+1}{2})}\right]}_{U|T=t) \text{ is } Gamma(\frac{d+1}{2},\lambda)} \end{aligned}$$

$$\Rightarrow \quad f_T(t) = \frac{1}{\sqrt{d}B\left(\frac{1}{2},\frac{d}{2}\right)}\left(1+\frac{t^2}{d}\right)^{-\frac{d+1}{2}}$$

F distribution F_{d_1,d_2}

Definition

$$\frac{V_1/d_1}{V_2/d_2} \qquad \text{where} \quad V_1 \sim \chi^2_{d_1} \text{ and } V_2 \sim \chi^2_{d_2} \text{ independent}$$

PDF

$$f_F(x) = \frac{1}{B\left(\frac{d_1}{2},\frac{d_2}{2}\right)x} \cdot \sqrt{\frac{(d_1x)^{d_1}\cdot d_2^{d_2}}{(d_1x+d_2)^{d_1+d_2}}} \qquad \text{for } x>0$$

With $f = \frac{x/d_1}{y/d_2}$ and $z = y$, where $X \sim \chi^2_{d_1}$ and $Y \sim \chi^2_{d_2}$ are independent,

$$\left|\frac{\partial(x,y)}{\partial(f,z)}\right| = \left|\frac{\partial(f,z)}{\partial(x,y)}\right|^{-1} = \left| det \begin{pmatrix} \frac{1/d_1}{z/d_2} & * \\ 0 & 1 \end{pmatrix}\right|^{-1} = \frac{z/d_2}{1/d_1}$$

With $\lambda = \frac{1}{2}(1+\frac{d_1}{d_2}f)$

$$\begin{aligned}
f_{F,Z}(f,z) &= f_{X,Y}(x,y)\left|\frac{\partial(x,y)}{\partial(f,z)}\right| \\
&= \frac{\frac{1}{2}\left(\frac{1}{2}x\right)^{\frac{d_1}{2}-1} e^{-\frac{1}{2}x}}{\Gamma\left(\frac{d_1}{2}\right)} \cdot \frac{\frac{1}{2}\left(\frac{1}{2}y\right)^{\frac{d_2}{2}-1} e^{-\frac{1}{2}y}}{\Gamma\left(\frac{d_2}{2}\right)} \cdot \frac{z/d_2}{1/d_1} \\
&= \frac{\frac{1}{2}\left(f\frac{d_1}{d_2}\right)^{\frac{d_1}{2}-1}}{\Gamma\left(\frac{d_1}{2}\right)} \cdot \frac{\left(\frac{1}{2}z\right)^{\frac{d_1+d_2}{2}-1} e^{-\frac{1}{2}(1+\frac{d_1}{d_2}f)z}}{\Gamma\left(\frac{d_2}{2}\right)} \cdot \frac{1/d_2}{1/d_1} \\
&= \frac{1}{B\left(\frac{d_1}{2},\frac{d_2}{2}\right)f} \cdot \sqrt{\frac{(d_1f)^{d_1}\cdot d_2^{d_2}}{(d_1f+d_2)^{d_1+d_2}}} \cdot \underbrace{\left[\frac{\lambda(\lambda z)^{\frac{d_1+d_2}{2}-1} e^{-\lambda z}}{\Gamma\left(\frac{d_1+d_2}{2}\right)}\right]}_{Z|F=f) \text{ is } Gamma(\frac{d_1+d_2}{2},\lambda)}
\end{aligned}$$

$$\Rightarrow \quad f_F(f) = \frac{1}{B\left(\frac{d_1}{2},\frac{d_2}{2}\right)f} \cdot \sqrt{\frac{(d_1f)^{d_1}\cdot d_2^{d_2}}{(d_1f+d_2)^{d_1+d_2}}} \qquad \text{for } f>0$$