Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 29 (2026), No. 5, pp. 487–515
DOI : 10.47974/JSMS-1554

A Bivariate mixture of student's *t* and normal distribution : Properties and estimation

G. S. David Sam Jayakumar *
Jamal Institute of Management
Jamal Mohamed College (Affiliated to Bharathidasan University)
Trichy 620020
Tamil Nadu
India

W. Samuel †
Department of Business Administration
Jamal Mohamed College (Affiliated to Bharathidasan University)
Trichy 620020
Tamil Nadu
India

Bejoy John Thomas §
Rajagiri College of Social Sciences
Rajagiri Business School
Kochi
Kerala
India

Abstract

This paper investigates a Bivariate mixture of the student's t and Normal distributions, referred to as the *t*-Normal distribution, where the marginals are univariate Student's *t* and Normal distributions, respectively. The study derives the conditional distributions along with their associated constants. It also presents several generating functions, including the moment, cumulant, characteristic, hazard, survival, cumulative hazard functions, and Shannon's differential entropy. Three-dimensional probability surfaces illustrate the

* *ORCID-ID:* **0000-0002-6630-8783**
† *ORCID-ID:* **0000-0002-9112-9744**

* *E-mail:* **samjaya77@gmail.com (Corresponding Author)**
† *E-mail:* **wsm@jmc.edu;**
wsamuel365@gmail.com
§ *E-mail:* **bejoy@rajagiri.edu**

shape of the *t*-Normal density. Special cases of this distribution include mixtures involving logarithmic, logit, and hyperbolic sine transformations. Furthermore, the paper derives the maximum likelihood estimators for the parameters, calculates the elements of the Fisher Information matrix and explores parameter estimation through a nonlinear programming approach.

Subject Classification: *62H10.*

Keywords: *Student's t distribution, Normal distribution, Bivariate mixture, Three-dimensional probability surfaces, Parameter estimation, Fisher information matrix, Non-linear programming.*

Some Preliminaries

The explicit form of the t-Normal distribution's PDF, as well as the computation of normalization constants, generating functions, and parameter estimates, involves various special functions, as outlined in Prudnikov, Brychkov, and Marichev [2] and Gradshteyn and Ryzhik [1]. These are presented as follows:

1. The Generalized hyper geometric function of order p and q is defined as

$$ {}_pF_q(a_1,a_2,\ldots,a_p;b_1,b_2,\ldots,b_q;x)=\sum_{k=0}^{\infty}\frac{(a_1)_k(a_2)_k\ldots(a_p)_k}{(b_1)_k(b_2)_k\ldots(b_q)_k}\left(\frac{x^k}{k!}\right)$$

2. The rising factorial or Pochammer symbol is given as

$$(a)_k=a(a+1)(a+2)\ldots\ldots(a+k-1)$$

3. The integral representation of the Lower Incomplete Gamma function as

$$\gamma(n,x)=\int_0^x t^{n-1}e^{-t}dt$$

4. The integral representation of the Beta function is given as

$$B(x;a,b)=\int_0^x\frac{t^{a-1}}{(1+t)^{a+b}}dt$$

5. The digamma function is defined as the logarithmic derivative of the Gamma function and is given by.

$$\Psi(x)=\Gamma'(x)/\Gamma(x)$$

Section 1 - Introduction

In The Seven Pillars of Statistics, Stigler [3] identified regression as the fifth pillar of statistical science, tracing its conceptual origins to Galton's 1885 work and its formal explanation through the Bivariate normal distribution. Beyond regression, Stigler [4] emphasized the binomial distribution as a historical cornerstone of statistics, noting that it emerged from both practical and philosophical challenges and facilitated statistical inference long before the development of modern data science. Building on these foundational ideas, Fisher [6] established Student's t distribution as a practical and indispensable tool for statistical inference, particularly in contexts involving small sample sizes and unknown population variance. This advancement was rooted in Gosset's seminal contribution [9], which fundamentally reshaped statistical inference by demonstrating that reliable estimation of a population mean from small samples requires a distribution that explicitly accounts for the uncertainty associated with variance estimation. Complementing these theoretical developments, McMullen's work [10], while not introducing new statistical theory, provides valuable insight into the character, motivations, and practical statistical philosophy of William Sealy Gosset—the man behind the pseudonym "Student."

The origins of the normal distribution can be traced back to Abraham De Moivre (1667–1754) [5], who first derived it as the limiting case of the binomial distribution. Building on this foundation, Carl Friedrich Gauss further developed the distribution around 1800 in his quest to establish the law of error in astronomical observations. Interestingly, the term "normal" did not appear in the literature until around 1870. Over the following century, as numerous probability distributions were introduced, the normal distribution emerged as a standard due to its flexibility and wide range of applications across scientific disciplines. In the context of sampling theory, the normal distribution is typically regarded as a large-sample or asymptotic distribution. In 1904, Gosset [7], writing under the pseudonym "Student," published an internal report titled *The Application of the Law of Error*. This work marked a significant step toward small-sample inference. Subsequent correspondence between Gosset and Ronald Fisher eventually led to the development of the student's t-distribution. Although Fisher initially resisted Karl Pearson's [8] introduction of the chi-square variate, he later formally investigated the t-distribution, publishing results in 1925 that confirmed its equivalence to Gosset's original formulation. In modeling complex data, mixtures of distributions such as the Bivariate Student's t and Normal are often employed to capture both heavy-tailed and light-tailed behavior. Since Shannon's entropy quantifies the uncertainty associated with a probability distribution, understanding its bounds becomes crucial in assessing the information content of such mixtures. Recently a new and strong bounds obtained for Shannon's entropy of a probability distribution by [11]. Today, the normal and Student's *t*-distributions are recognized in statistical literature as the foundational population and sampling distributions, respectively. This dual role inspired the authors to explore a synthesis of the two, resulting in a Bivariate mixture of the normal and Student's *t*-distributions—termed the *t*-Normal distribution. The subsequent sections of this work delve into the structural properties, statistical characteristics, and parameter estimation techniques associated with the *t*-Normal distribution.

Section 2 - Bivariate mixture of student's t-normal distribution

Definition 2.1: Let X and Y be the jointly distributed random variables follows Bivariate mixture of t-normal distribution with 'v' degrees of freedom, then the density function of the distribution is defined as

$$f_{X,Y}(x,y;\mu_X,\mu_Y,\sigma_X,\sigma_Y,v)=\frac{(v/2)^{v/2}}{\sigma_X\sigma_Y\,\Gamma(v/2)\sqrt{2\pi}}$$

$$\left|\frac{x-\mu_X}{\sigma_X}\right|^{v}\left|\frac{y-\mu_Y}{\sigma_Y}\right|^{-(v+1)}\exp\left(-\frac{1}{2}\left(1+\frac{1}{\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2/v}\right)\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right) \tag{2.1}$$

Where $-\infty<x,y<+\infty,-\infty<\mu_X,\mu_Y<+\infty,\sigma_X,\sigma_Y,v>0$

Result 2.2: From (2.1), if $z_X=(x-\mu_X)/\sigma_X$ and $z_Y=(y-\mu_Y)/\sigma_Y$, then the density function of Bivariate mixture of standard t-normal distribution with degrees of freedom' v' is given as

$$f_{Z_X,Z_Y}(z_X,z_Y;v)=\frac{(v/2)^{v/2}}{\Gamma(v/2)\sqrt{2\pi}}\left|z_X\right|^{v}\left|z_Y\right|^{-(v+1)}\exp\left(-\frac{1}{2}\left(1+\frac{1}{z_Y^2/v}\right)z_X^2\right) \tag{2.2}$$

Where $-\infty<z_X,z_Y<+\infty,\ v>0$

Result 2.3: The following probability surfaces of Bivariate mixture of standard t-normal distribution for the selected values of 'v' are visualized below:

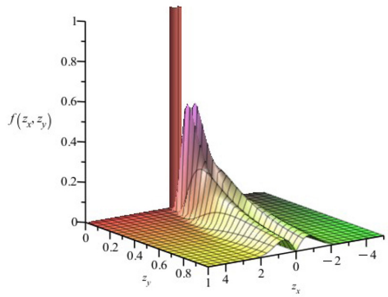

Figure 1

v = 1

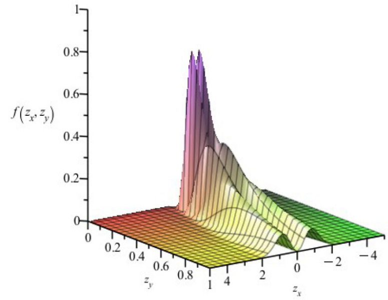

Figure 2

v = 2

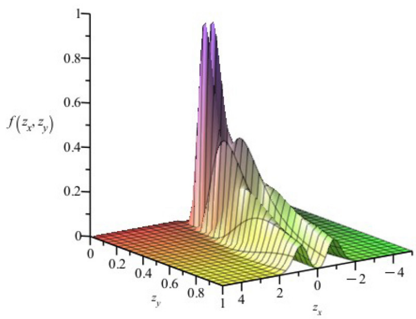

Figure 3

v = 3

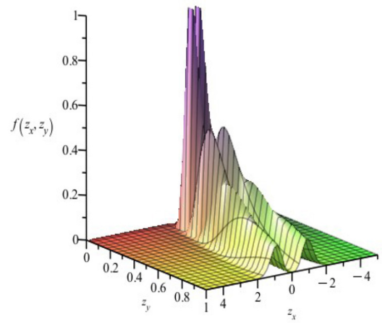

Figure 4

v = 4

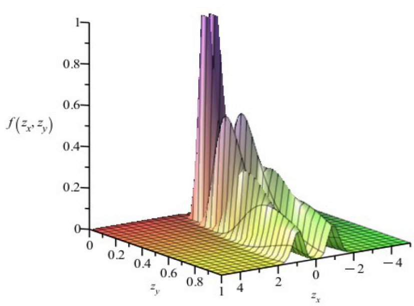

Figure 5

v = 5

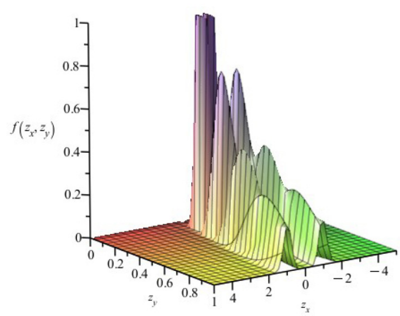

Figure 6

v = 10

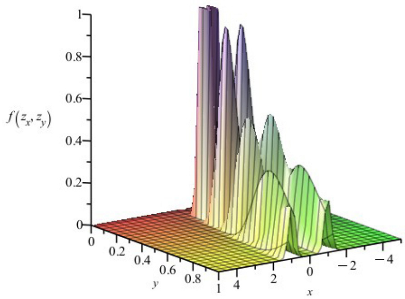

Figure 7

v = 15

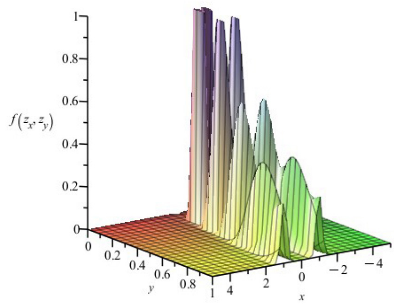

Figure 8

v = 20

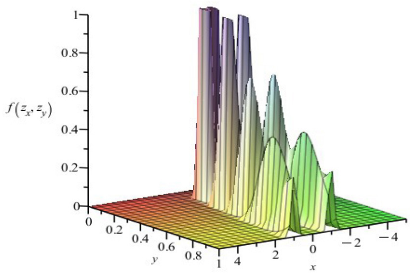

Figure 9

v = 25

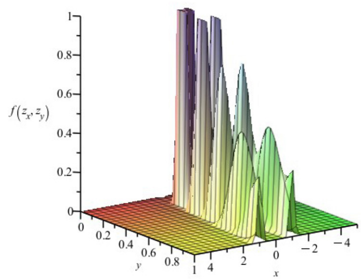

Figure 10

v = 30

Theorem 2.3: *The cdf of the Bivariate-scaled t-normal distribution is obtained as*

$$F_{X,Y}(x,y)=\frac{1}{4\Gamma(v/2)\sqrt{\pi}}$$

$$\left(B\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2/v;1/2,v/2\right)\gamma\left(\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2;(v+1)/2\right)-\sqrt{\pi}\,\Gamma(v/2)\right) \quad (2.3)$$

where $B(\ )$ *and* $\gamma(\ )$ *are the Incomplete beta and lower incomplete Gamma function respectively.*

Proof: Let the cumulative distribution function of a Bivariate distribution is

$$F_{X,Y}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f_{U,S}(u,s)\,du\,ds$$

$$= \frac{(v/2)^{v/2}}{\sigma_X \sigma_Y \, \Gamma(v/2)\sqrt{2\pi}} \int_{-\infty}^{x}\int_{-\infty}^{y} \left|\frac{u-\mu_X}{\sigma_X}\right|^{v} \left|\frac{s-\mu_Y}{\sigma_Y}\right|^{-(v+1)}$$

$$\exp\left(-\frac{1}{2}\left(1+\frac{1}{\left(\frac{s-\mu_Y}{\sigma_Y}\right)^2 / v}\right)\left(\frac{u-\mu_X}{\sigma_X}\right)^2\right) duds \quad (2.4)$$

Now rewrite (2.4) in terms of standardized form, then the Integral becomes

$$= \int_{-\infty}^{\frac{x-\mu_X}{\sigma_X}} \int_{-\infty}^{\frac{y-\mu_Y}{\sigma_Y}} \frac{(v/2)^{v/2}}{\Gamma(v/2)\sqrt{2\pi}} |U|^{v} |S|^{-(v+1)} \exp\left(-\frac{1}{2}\left(1+\frac{1}{S^2/v}\right)U^2\right) dUdS \quad (2.5)$$

Split the integrals into additive form in (2.5) and perform the integration, then the final expression of CDF as

$$F_{X,Y}(x,y) = \frac{1}{4\,\Gamma(v/2)\sqrt{\pi}}$$

$$\left(B\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2 / v; 1/2, v/2\right)\gamma\left(\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2; (v+1)/2\right) - \sqrt{\pi}\,\Gamma(v/2)\right)$$

Section 3 - Marginal, Conditional distributions and Conditional Moments

Theorem 3.1: *The marginal distribution of X for (2.1) is the Scaled Gaussian normal distribution and it's density function is given as*

$$f_X(x;\mu_X,\sigma_X) = \frac{1}{\sigma_X\sqrt{2\pi}} \exp\left(\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right) \quad (3.1)$$

where $-\infty < x < +\infty,\ -\infty < \mu_X < +\infty, \sigma_X > 0$

Proof: The marginal distribution of *X* is derived as

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\, dy$$

$$= \frac{(v/2)^{v/2}}{\sigma_X \sigma_Y \, \Gamma(v/2)\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left| \frac{x-\mu_X}{\sigma_X} \right|^{v} \left| \frac{y-\mu_Y}{\sigma_Y} \right|^{-(v+1)} \exp\left(-\frac{1}{2} \left(1 + \frac{1}{\left(\frac{y-\mu_Y}{\sigma_Y} \right)^2 / v} \right) \left(\frac{x-\mu_X}{\sigma_X} \right)^2 \right) dy$$

By setting $z_Y = (y-\mu_Y)/\sigma_Y$ and perform the integration, then the final result is found to be

$$f_X(x;\mu_X,\sigma_X) = \frac{1}{\sigma_X\sqrt{2\pi}} \exp\left(\frac{1}{2} \left(\frac{x-\mu_X}{\sigma_X} \right)^2 \right)$$

where $-\infty < x < +\infty,\ -\infty < \mu_X < +\infty, \sigma_X > 0$

Theorem 3.2: *The marginal distribution of Y for (2.1) is the Scaled student's t distribution with 'v' degrees of freedom and it's density function is given as*

$$f_Y(y;\mu_Y,\sigma_Y) = \frac{1}{\sigma_Y \sqrt{v} B(v/2, 1/2)} \left(1 + \left(\frac{y-\mu_Y}{\sigma_Y} \right)^2 / v \right)^{-\frac{v+1}{2}} \tag{3.2}$$

Where $-\infty < y < +\infty, -\infty < \mu_Y < +\infty, \sigma_Y, v > 0$

Proof: The marginal distribution of Y is computed as

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\, dx$$

$$= \frac{(v/2)^{v/2}}{\sigma_X \sigma_Y \, \Gamma(v/2)\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left| \frac{x-\mu_X}{\sigma_X} \right|^{v} \left| \frac{y-\mu_Y}{\sigma_Y} \right|^{-(v+1)} \exp\left(-\frac{1}{2} \left(1 + \frac{1}{\left(\frac{y-\mu_Y}{\sigma_Y} \right)^2 / v} \right) \left(\frac{x-\mu_X}{\sigma_X} \right)^2 \right) dy$$

By Setting $z_X = (x-\mu_X)/\sigma_X$, then the result is found to be

$$f_Y(y;\mu_Y,\sigma_Y) = \frac{1}{\sigma_Y\sqrt{v}B(v/2,1/2)}\left(1+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2/v\right)^{-\frac{v+1}{2}}$$

Where $-\infty < y < +\infty, -\infty < \mu_Y < +\infty, \sigma_Y, v > 0$

Theorem 3.3: *The Probability density function of Conditional Bivariate mixture of scaled t-normal distribution of X on Y is*

$$f_{X/Y}(x/y) = \frac{1}{\sigma_X\,\Gamma\left(\frac{v+1}{2}\right)}\left(\frac{1}{2}\left(1+\frac{1}{\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2/v}\right)\right)^{\frac{v+1}{2}}\left|\frac{x-\mu_X}{\sigma_X}\right|^v$$
$$\exp\left(-\frac{1}{2}\left(1+\frac{1}{\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2/v}\right)\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right) \qquad (3.3)$$

where $-\infty < x < +\infty,\ -\infty < \mu_X < +\infty, \sigma_X > 0$

Proof: It is obtained from $f_{X/Y}(x/y) = f_{X,Y}(x,y)/f_Y(y)$

Theorem 3.4: *The r^{th} order moment of the Bivariate mixture of Conditional t-normal distribution of X on Y is given as*

$$E_{X/Y}(x^r/y) = \frac{\mu_X^r}{\Gamma\left(\frac{v+1}{2}\right)}\sum_{q=0}^{r}\left[\binom{r}{2q}(\sqrt{2}\,\sigma_X/\mu_X)^{2q}\,\Gamma\left(\frac{v+1}{2}+q\right)\left(1+\frac{1}{\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2/v}\right)^{-q}\right]$$

(3.4)

Proof: The r^{th} order moment of a conditional distribution is

$$E_{X/Y}(x^r / y) = \int_{-\infty}^{+\infty} x^r f_{X/Y}(x/y)\, dx$$

$$= \frac{1}{\sigma_X \, \Gamma\left(\frac{v+1}{2}\right)} \left(\frac{1}{2} \left(1 + \frac{1}{\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 / v} \right) \right)^{\frac{v+1}{2}} \int_{-\infty}^{+\infty} x^r \left| \frac{x-\mu_X}{\sigma_X} \right|^v \exp\left(-\frac{1}{2} \left(1 + \frac{1}{\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 / v} \right) \left(\frac{x-\mu_X}{\sigma_X} \right)^2 \right) dx \tag{3.5}$$

Then set $z_X = (x - \mu_X)/\sigma_X$ in (3.5), then the integral becomes

$$= \int_{-\infty}^{+\infty} \frac{1}{\Gamma\left(\frac{v+1}{2}\right)} \left(\frac{1}{2} \left(1 + \frac{1}{\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 / v} \right) \right)^{\frac{v+1}{2}} \int_{-\infty}^{+\infty} \left(\mu_X + \sigma_X z_X \right)^r \left| z_X \right|^v \exp\left(-\frac{1}{2} \left(1 + \frac{1}{\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 / v} \right) z_X^2 \right) dz_X \tag{3.6}$$

From (3.6), expand the binomial series and perform the integration, the result follows on writing

$$E_{X/Y}(x^r / y) = \frac{\mu_X^r}{\Gamma\left(\frac{v+1}{2}\right)} \sum_{q=0}^{r} \left[\binom{r}{2q} (\sqrt{2}\, \sigma_X / \mu_X)^{2q} \, \Gamma\left(\frac{v+1}{2} + q \right) \left(1 + \frac{1}{\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 / v} \right)^{-q} \right] \tag{3.7}$$

From (3.7), if r = 1, 2, 3, 4 then the first and higher order conditional moment is

$$E_{X/Y}(x/y) = \mu_X$$

$$E_{X/Y}(x^2 / y) = \mu_X^2 \left(1 + (v+1)(\sigma_X / \mu_X)^2 \left(1 + \left(\frac{y-\mu_Y}{\sigma_Y} \right)^2 / v \right) \right)$$

$$E_{X/Y}(x^3 / y) = \mu_X^3 \left(1 + 3(v+1)(\sigma_X / \mu_X)^2 \left(1 + \left(\frac{y-\mu_Y}{\sigma_Y} \right)^2 / v \right) \right)$$

$$E_{X/Y}(x^4/y)=\mu_X^4\left(1+6(v+1)(\sigma_X/\mu_X)^2\left(1+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2/v\right)\right.$$
$$\left.+(v+1)(v+3)(\sigma_X/\mu_X)^4\left(1+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2/v\right)^2\right)$$

Theorem 3.5: *The pdf of Bivariate mixture of Conditional t-normal distribution of Y on X is*

$$f_{Y/X}(y/x)=\frac{(v/2)^{v/2}}{\sigma_Y\ \Gamma(v/2)}\left|\frac{x-\mu_X}{\sigma_X}\right|^v\left|\frac{y-\mu_Y}{\sigma_Y}\right|^{-(v+1)}\exp\left(-\frac{1}{2}\left(\frac{\left(\frac{x-\mu_X}{\sigma_X}\right)^2}{\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2/v}\right)\right)\quad(3.8)$$

Where $-\infty<y<+\infty,-\infty<\mu_Y<+\infty,\sigma_Y,v>0$

Proof: It is obtained from $f_{Y/X}(y/x)=f_{X,Y}(x,y)/f_X(x)$

Theorem 3.6: *The r^{th} conditional moment of the distribution of Y on X is*

$$E_{Y/X}(y^r/x)=\frac{(\mu_Y)^r}{\Gamma(v/2)}\sum_{q=0}^{r}\left[\binom{r}{2q}(v/2)^{q/2}(\sigma_Y/\mu_Y)^{2q}\left(\frac{x-\mu_X}{\sigma_X}\right)^q\Gamma\left(\frac{v-q}{2}\right)\right]\quad(3.9)$$

where $\Gamma(\ )$ is the Gamma function.

Proof: The r^{th} order moment of a distribution is

$$E_{Y/X}(y^r/x)=\int_{-\infty}^{+\infty}y^r f_{Y/X}(y/x)\,dy$$

$$=\frac{(v/2)^{v/2}}{\sigma_Y\ \Gamma(v/2)}\left|\frac{x-\mu_X}{\sigma_Y}\right|^v\int_{-\infty}^{+\infty}y^r\left|\frac{y-\mu_X}{\sigma_Y}\right|^{-(v+1)}\exp\left(-\frac{1}{2}\left(\frac{\left(\frac{x-\mu_X}{\sigma_Y}\right)^2}{\left(\frac{y-\mu_X}{\sigma_Y}\right)^2/v}\right)\right)dy\quad(3.10)$$

Then set $z_Y=(y-\mu_Y)/\sigma_Y$ in (3.10), then the integral becomes

$$E_{Y/X}(y^r/x) == \frac{(v/2)^{v/2}}{\Gamma(v/2)}\left|\frac{x-\mu_X}{\sigma_X}\right|^v \int_{-\infty}^{+\infty}(\mu_Y+\sigma_Y z_Y)^r \left|z_Y\right|^{-(v+1)} \exp\left(-\frac{1}{2}\left(\frac{\left(\frac{x-\mu_X}{\sigma_X}\right)^2}{(z_Y)^2/v}\right)\right) dz_Y \tag{3.11}$$

From (3.11), expand the binomial series and perform the integration, the result is found to be

$$E_{Y/X}(y^r/x) = \frac{(\mu_Y)^r}{\Gamma(v/2)}\sum_{q=0}^{r}\left[\binom{r}{2q}(v/2)^{q/2}(\sigma_Y/\mu_Y)^{2q}\Gamma\left(\frac{v-q}{2}\right)\left(\frac{x-\mu_X}{\sigma_X}\right)^q\right] \tag{3.12}$$

From (3.12) if r=1, 2, 3, 4 then the first and high order conditional moments are

$$E_{Y/X}(y/x) = \mu_Y$$

$$E_{Y/X}(y^2/x) = \mu_Y^2\left(1+(\sigma_Y/\mu_Y)^2\frac{\Gamma((v/2)-(1/2))}{\Gamma(v/2)}\left(\frac{x-\mu_X}{\sigma_X}\right)\right)$$

$$E_{Y/X}(y^3/x) = \mu_Y^3\left(1+\sqrt{3}(\sigma_Y/\mu_Y)^3\frac{\Gamma((v/2)-(1/2))}{\Gamma(v/2)}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right)$$

$$E_{Y/X}(y^4/x) = \mu_Y^4\left(1+\sqrt{6}(\sigma_Y/\mu_Y)^2\frac{\Gamma((v/2)-(1/2))}{\Gamma(v/2)}\left(\frac{x-\mu_X}{\sigma_X}\right)\right.$$
$$\left.+(\sigma_Y/\mu_Y)^4\frac{\Gamma((v/2)-1)}{\Gamma(v/2)}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right)$$

Section 4 - Constants and Generating functions

Theorem 4.1: *If X and Y are jointly distributed according to Equation (2.1), then the product moment is.*

$$E_{XY}(xy) = \mu_X\mu_Y \tag{4.1}$$

Proof: The result is obtained by writing.

$$E_{XY}(xy) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy\ f_{X,Y}(x,y)\,dxdy$$

$$= \frac{(v/2)^{v/2}}{\sigma_X\sigma_Y\ \Gamma(v/2)\sqrt{2\pi}} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy \left|\frac{x-\mu_X}{\sigma_X}\right|^{v} \left|\frac{y-\mu_Y}{\sigma_Y}\right|^{-(v+1)} \exp\left(-\frac{1}{2}\left(1+\frac{1}{\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2/v}\right)\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right)dxdy$$

(4.2)

By setting $z_X = (x-\mu_X)/\sigma_X$ and $z_Y = (y-\mu_Y)/\sigma_Y$ in (4.2), then the integral becomes

$$= \frac{(v/2)^{v/2}}{\Gamma(v/2)\sqrt{2\pi}} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (\mu_X+\sigma_X z_X)(\mu_Y+\sigma_Y z_Y)\,|z_X|^{v}\,|z_Y|^{-(v+1)} \exp\left(-\frac{1}{2}\left(1+\frac{1}{z_Y^2/v}\right)z_X^2\right)dz_X dz_Y$$

(4.3)

From (4.3), perform the integration, then the first order product moment found to be

$$E_{XY}(xy) = \mu_X\mu_Y$$

Corollary 1: *If X and Y are jointly distributed according to (2.1) then the Product moment, Co-variance and Correlation Co-efficient between the mixture of t-normal variables are given as*

$$E_{XY}(xy) = \mu_X\mu_Y \tag{4.4}$$

$$COV_{X,Y}(x,y) = 0 \tag{4.5}$$

$$\rho_{X,Y}(x,y) = 0 \tag{4.6}$$

Therefore (4.5) and (4.6) clearly visualizes the mixture of t-normal variables are uncorrelated but not independent and the uncorrelatedness does not implies independence. Hence, the proposed density function of t-normal mixture is unique and it's having several notable properties.

Theorem 4.2: *The Joint Shannon's differential entropy of the Bivariate mixture of t-normal distribution is*

$$h'_{X,Y}(x,y) = -(\omega_1(\sigma_X,\sigma_Y,v)+\omega_2(v)+\omega_3(v)+\omega_4(v)) \tag{4.7}$$

where $$\omega_1(\sigma_X, \sigma_Y, v) = -\log\left(\frac{(v/2)^{v/2}}{\sqrt{2\pi}\sigma_x\sigma_y\ \Gamma(v/2)}\right)$$

$$\omega_2(v) = \frac{-v/2}{\Gamma(v/2)\sqrt{\pi}} B(v/2, 1/2)(\Gamma((v+1)/2)(\Psi((v+1)/2) - \log 2) - \pi\tan(\pi v/2) + v\,{}_3F_2(1,1,1+v/2;3/2,2;1))$$

$$\omega_3(v) = \frac{v+1}{2}\left(\log(\sqrt{v}) - \frac{1}{2}(\Psi(v/2) + \gamma + 2\log 2)\right)$$

$$\omega_4(v) = \frac{B(v/2, 1/2)}{4B(v/2, 3/2)}$$

and ${}_3F_2(\ )$, $\Psi(\ )$, $B(\ )$ are the Generalized hypergeometric,di-gamma and Beta functions respectively.

Proof: It is found from

$$h'_{X,Y}(x,y) = -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\log(f_{X,Y}(x,y))\,dxdy$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(\begin{array}{l} -\log\left(\dfrac{(v/2)^{v/2}}{\sigma_x\sigma_y\ \Gamma(v/2)\sqrt{2\pi}}\left|\dfrac{x-\mu_x}{\sigma_x}\right|^{v}\left|\dfrac{y-\mu_y}{\sigma_y}\right|^{-(v+1)} e^{-\frac{1}{2}\left(1+\frac{1}{\left(\left(\frac{y-\mu_y}{\sigma_y}\right)^2/v\right)}\right)\left(\frac{x-\mu_x}{\sigma_x}\right)^2}\right) \\ \times\dfrac{(v/2)^{v/2}}{\sigma_x\sigma_y\ \Gamma(v/2)\sqrt{2\pi}}\left|\dfrac{x-\mu_x}{\sigma_x}\right|^{v}\left|\dfrac{y-\mu_y}{\sigma_y}\right|^{-(v+1)} e^{-\frac{1}{2}\left(1+\frac{1}{\left(\left(\frac{y-\mu_y}{\sigma_y}\right)^2/v\right)}\right)\left(\frac{x-\mu_x}{\sigma_x}\right)^2} \end{array}\right) dxdy$$

By integration, the final result will be

$$h'_{X,Y}(x,y) = -(\omega_1(\sigma_X, \sigma_Y, v) + \omega_2(v) + \omega_3(v) + \omega_4(v))$$

Theorem 4.3: *The Moment generating function of the Bivariate t-normal mixture is*

$$M_{X,Y}(t_1,t_2)=\frac{(1/2)^{v/2}e^{\mu_X t_1+\mu_Y t_2}}{\Gamma(v/2)\sqrt{2\pi}}$$
$$\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\left[(\sqrt{2}\mu_X)^{2k}(\sqrt{v}\mu_Y)^{2l}\,\Gamma(k+l+(1/2))\Gamma((v/2)-l)\frac{t_1^{2k}t_2^{2l}}{(2k)!(2l)!}\right]$$

Proof: Let the moment generating function of the Bivariate distribution be given as.

$$M_{X,Y}(t_1,t_2)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}e^{t_1x+t_2y}\left|\frac{x-\mu_X}{\sigma_X}\right|^{v}\left|\frac{y-\mu_Y}{\sigma_Y}\right|^{-(v+1)}\exp\left(-\frac{1}{2}\left(1+\frac{1}{\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2/v}\right)\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right)dxdy \tag{4.8}$$

By setting $z_X=(x-\mu_X)/\sigma_X$, $z_Y=(y-\mu_Y)/\sigma_Y$ and expand the exponent into the sum of odd and even power series as in (4.8) and perform the integration, then the odd powers are vanished, finally the result is found to be

$$M_{X,Y}(t_1,t_2)=\frac{(1/2)^{v/2}e^{\mu_X t_1+\mu_Y t_2}}{\Gamma(v/2)\sqrt{2\pi}}$$
$$\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\left[(\sqrt{2}\mu_X)^{2k}(\sqrt{v}\mu_Y)^{2l}\,\Gamma(k+l+(1/2))\Gamma((v/2)-l)\frac{t_1^{2k}t_2^{2l}}{(2k)!(2l)!}\right]$$

Theorem 4.4: *The Cumulant of the Bivariate mixture of t-normal distribution is*

$$C_{X,Y}(t_1,t_2)=\log\left(\frac{(1/2)^{v/2}e^{\mu_X t_1+\mu_Y t_2}}{\Gamma(v/2)\sqrt{2\pi}}\right)+$$
$$\log\left(\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\left[(\sqrt{2}\mu_X)^{2k}(\sqrt{v}\mu_Y)^{2l}\,\Gamma(k+l+(1/2))\Gamma((v/2)-l)\frac{t_1^{2k}t_2^{2l}}{(2k)!(2l)!}\right]\right) \tag{4.9}$$

Proof: It is found from $C_{X,Y}(t_1,t_2)=\log M_{X,Y}(t_1,t_2)$

Theorem 4.5: *The Characteristic function of the Bivariate t-normal mixture is*

$$\phi_{X,Y}(t_1,t_2)=\frac{(1/2)^{v/2}e^{i(\mu_X t_1+\mu_Y t_2)}}{\Gamma(v/2)\sqrt{2\pi}}$$
$$\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\left[(\sqrt{2}\mu_X)^{2k}(\sqrt{v}\mu_Y)^{2l}\Gamma(k+l+(1/2))\Gamma((v/2)-l)\frac{t_1^{2k}t_2^{2l}}{(2k)!(2l)!}\right] \quad (4.10)$$

Proof: Let the Characteristic function of a Bivariate distribution is given as

$$\phi_{X,Y}(t_1,t_2)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}e^{i(t_1x+t_2y)}\left|\frac{x-\mu_X}{\sigma_X}\right|^{v}\left|\frac{y-\mu_Y}{\sigma_Y}\right|^{-(v+1)}\exp\left(-\frac{1}{2}\left(1+\frac{1}{\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2/v}\right)\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right)dxdy$$

From Theorem 3.2, the integration is straightforward, and the final form of the characteristic function is obtained as.

$$\phi_{X,Y}(t_1,t_2)=\frac{(1/2)^{v/2}e^{i(\mu_X t_1+\mu_Y t_2)}}{\Gamma(v/2)\sqrt{2\pi}}$$
$$\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\left[(\sqrt{2}\mu_X)^{2k}(\sqrt{v}\mu_Y)^{2l}\Gamma(k+l+(1/2))\Gamma((v/2)-l)\frac{t_1^{2k}t_2^{2l}}{(2k)!(2l)!}\right]$$

Result 4.6: Table-1 shows the special cases of (2.1) based on Jacobian transformation of t-normal vatiables.

Table-1

Special Cases

Case No.	**Distributions**	**Transformation**	
		X	**Y**
1	**Log (t-normal)**	e^x	e^y
2	**Logit (t-normal)**	$e^x/(1+e^x)$	$e^y/(1+e^y)$
3	**t-log normal**	**x**	e^y
4	**Log t-normal**	e^x	**y**
5	**t-logit normal**	**x**	$e^y/(1+e^y)$
6	**Logit t-normal**	$e^x/(1+e^x)$	**y**

Contd...

7	**Log t-Logit normal**	e^x	$e^y/(1+e^y)$
8	**Logit t-Log normal**	$e^x/(1+e^x)$	e^y
9	**Sine hyperbolic normal –t distribution**	$\text{Sinh}(x)$	y
10	**normal – Sine hyperbolic t**	x	$\text{Sinh}(y)$
11	**Sine hyperbolic (t-normal**	$\text{Sinh}(x)$	$\text{Sinh}(y)$
12	**Log normal – Sine hyperbolic t**	e^x	$\text{Sinh}(y)$
13	**Sine hyperbolic normal-log t**	$\text{Sinh}(x)$	e^x
14	**Logit normal-Sine hyperbolic t**	$e^x/(1+e^x)$	$\text{Sinh}(y)$
15	**Sine hyperbolic normal-logit t**	$\text{Sinh}(x)$	$e^y/(1+e^y)$

Section 5 - Parameter Estimation

Result 5.1: As an initial step, we apply estimation using the method of maximum likelihood. The log-likelihood function for a Bivariate random sample is. $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots, (x_{n-1}, y_{n-1}), (x_n, y_n)$ from (2.1)

$$\log L(x, y; \mu_X, \mu_Y, \sigma_X, \sigma_Y, v) = \left\{ \begin{array}{l} -\frac{n}{2}\log(2\pi) + \frac{nv}{2}\log(v/2) - n\log(\Gamma(v/2)) - n\log(\sigma_X \sigma_Y) + v\sum_{i=1}^{n} \log\left|\frac{x_i - \mu_X}{\sigma_X}\right| \\ -(v+1)\sum_{i=1}^{n}\log\left|\frac{y_i - \mu_Y}{\sigma_Y}\right| - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i - \mu_X}{\sigma_X}\right)^2 - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\left(\frac{x_i - \mu_X}{\sigma_X}\right)^2}{\left(\frac{y_i - \mu_Y}{\sigma_Y}\right)^2 / v}\right) \end{array} \right\} \tag{5.1}$$

The first-order derivatives of Equation (5.1) with respect to the five parametersare.

$$\frac{\partial \log L}{\partial \mu_X} = v\sum_{i=1}^{n}\left(\frac{abs(1, -(x_i - \mu_X)/\sigma_X)}{|x_i - \mu_X|}\right) + \sum_{i=1}^{n}\left(\frac{x_i - \mu_X}{\sigma_X}\right) + \sum_{i=1}^{n}\left(\frac{\left(\frac{x_i - \mu_X}{\sigma_X}\right)}{\left(\frac{y_i - \mu_Y}{\sigma_Y}\right)^2 / v}\right) \tag{5.2}$$

$$\frac{\partial \log L}{\partial \mu_Y} = -(v+1)\sum_{i=1}^{n}\left(\frac{abs(1,-(y_i-\mu_Y)/\sigma_Y)}{|y_i-\mu_Y|}\right) - v\sum_{i=1}^{n}\left(\frac{\sigma_Y\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^3}\right) \quad (5.3)$$

$$\frac{\partial \log L}{\partial \sigma_X} = -\frac{n}{2\sigma_X} + (v+1)\sum_{i=1}^{n}\left(\left(\frac{x_i-\mu_X}{\sigma_X}\right)\frac{abs(1,-(x_i-\mu_X)/\sigma_X)}{|x_i-\mu_X|}\right) - \frac{1}{\sigma_X}\sum_{i=1}^{n}\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2 + \sum_{i=1}^{n}\left(\frac{\frac{1}{\sigma_X}\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^2/v}\right) \quad (5.4)$$

$$\frac{\partial \log L}{\partial \sigma_Y} = -\frac{n}{2\sigma_Y} + (v+1)\sum_{i=1}^{n}\left(\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)\frac{abs(1,-(x_i-\mu_Y)/\sigma_Y)}{|x_i-\mu_Y|}\right) - \sum_{i=1}^{n}\left(\frac{\sigma_Y\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^2/v}\right) \quad (5.5)$$

$$\frac{\partial \log L}{\partial v} = \frac{n}{2}\log(v/2) + n - \frac{n}{2}\Psi(v/2) + \sum_{i=1}^{n}\log\left|\frac{x_i-\mu_X}{\sigma_X}\right| - \sum_{i=1}^{n}\log\left|\frac{y_i-\mu_Y}{\sigma_Y}\right| - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^2}\right) \quad (5.6)$$

By setting these expressions equal to zero and solving them together, the authors obtain the maximum likelihood estimates of the five parameters. In order to construct the Fisher Information matrix, one requires the Second-order and Cross Partial derivatives of (5.1) are

$$\frac{\partial^2 \log L}{\partial \mu_X^2} = \frac{v}{\sigma_X}\sum_{i=1}^{n}\left(\frac{signum(1,-(x_i-\mu_X)/\sigma_X)}{|x_i-\mu_X|} - \frac{abs(1,-(x_i-\mu_X)/\sigma_X)^2}{|x_i-\mu_X|^2}\right) - \frac{n}{\sigma_X} + \frac{n}{\sigma_X}\sum_{i=1}^{n}\left(\frac{1}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^2/v}\right) \quad (5.7)$$

$$\frac{\partial^2 \log L}{\partial \mu_Y^2} = -\frac{v+1}{\sigma_Y}\sum_{i=1}^{n}\left(\frac{signum(1,-(y_i-\mu_Y)/\sigma_Y)}{\left|y_i-\mu_Y\right|} - \frac{abs(1,-(y_i-\mu_Y)/\sigma_Y)^2}{\left|y_i-\mu_Y\right|}\right) - 3v\sum_{i=1}^{n}\left(\frac{\sigma_Y\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^4}\right) \tag{5.8}$$

$$\frac{\partial^2 \log L}{\partial \sigma_X^2} = \left(\begin{array}{l} \frac{n}{2\sigma_X^2} + (v+1)\sum_{i=1}^{n}\left(\left(\frac{\left|x_i-\mu_X\right|}{\sigma_X^3}\right) signum(1,-(x_i-\mu_X)/\sigma_X) + \left(\frac{x_i-\mu_X}{\sigma_X^2}\right)\frac{abs(1,-(x_i-\mu_X)/\sigma_X)}{\left|x_i-\mu_X\right|}\right) \\ +\frac{4}{\sigma_X^3}\sum_{i=1}^{n}\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2 - 4\sum_{i=1}^{n}\left(\frac{\frac{1}{\sigma_X^3}\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^2/v}\right) \end{array}\right) \tag{5.9}$$

$$\frac{\partial^2 \log L}{\partial \sigma_Y^2} = \left(\begin{array}{l} \frac{n}{2\sigma_Y^2} + (v+1)\sum_{i=1}^{n}\left(\left(\frac{\left|y_i-\mu_Y\right|}{\sigma_Y^3}\right)\frac{signum(1,-(y_i-\mu_Y)/\sigma_Y)}{\left|y_i-\mu_Y\right|} + \left(\frac{y_i-\mu_Y}{\sigma_Y^3}\right)\frac{abs(1,-(y_i-\mu_Y)/\sigma_Y)}{\left|y_i-\mu_Y\right|}\right) \\ -\sum_{i=1}^{n}\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^2/v}\right) \end{array}\right) \tag{5.10}$$

$$\frac{\partial^2 \log L}{\partial v^2} = \frac{n}{v} - \frac{n}{2}\Psi(1, v/2) \tag{5.11}$$

$$\frac{\partial^2 \log L}{\partial \mu_X \partial \mu_Y} = \frac{1}{\sigma_Y}\sum_{i=1}^{n}\left(\frac{2\left(\frac{x_i-\mu_X}{\sigma_X}\right)}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^3/v}\right) \tag{5.12}$$

$$\frac{\partial^2 \log L}{\partial\mu_X \partial\sigma_X} = v\sum_{i=1}^{n}\left(\left(\frac{x_i-\mu_X}{\sigma_X}\right)\frac{signum(1,-(x_i-\mu_X)/\sigma_X)}{\sigma_X\left|x_i-\mu_X\right|}\right) - \sum_{i=1}^{n}\left(\frac{x_i-\mu_X}{\sigma_X^2}\right) - \sum_{i=1}^{n}\left(\frac{\dfrac{x_i-\mu_X}{\sigma_X^2}}{\left(\dfrac{y_i-\mu_Y}{\sigma_Y}\right)^2/v}\right) \tag{5.13}$$

$$\frac{\partial^2 \log L}{\partial\mu_X \partial\sigma_Y} = \frac{1}{\sigma_Y}\sum_{i=1}^{n}\left(\frac{2\left(\dfrac{x_i-\mu_X}{\sigma_X}\right)}{\left(\dfrac{y_i-\mu_Y}{\sigma_Y}\right)^3/v}\right) \tag{5.14}$$

$$\frac{\partial^2 \log L}{\partial\mu_X \partial v} = \sum_{i=1}^{n}\left(\frac{abs(1,-(x_i-\mu_X)/\sigma_X)}{\left|x_i-\mu_X\right|}\right) + \sum_{i=1}^{n}\left(\frac{x_i-\mu_X}{\sigma_X}\right) + \sum_{i=1}^{n}\left(\frac{\left(\dfrac{x_i-\mu_X}{\sigma_X}\right)}{\left(\dfrac{y_i-\mu_Y}{\sigma_Y}\right)^2}\right) \tag{5.15}$$

$$\frac{\partial^2 \log L}{\partial\mu_Y \partial\mu_X} = \frac{1}{\sigma_Y}\sum_{i=1}^{n}\left(\frac{2\left(\dfrac{x_i-\mu_X}{\sigma_X}\right)}{\left(\dfrac{y_i-\mu_Y}{\sigma_Y}\right)^3/v}\right) \tag{5.16}$$

$$\frac{\partial^2 \log L}{\partial\mu_Y \partial\sigma_X} = v\sum_{i=1}^{n}\left(\frac{\dfrac{2\sigma_Y}{\sigma_X}\left(\dfrac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^3}\right) \tag{5.17}$$

$$\frac{\partial^2 \log L}{\partial\mu_Y \partial\sigma_Y} = -(v+1)\sum_{i=1}^{n}\left(\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)\frac{signum(1,-(y_i-\mu_Y)/\sigma_Y)}{\sigma_Y\left|y_i-\mu_Y\right|}\right) - v\sum_{i=1}^{n}\left(\frac{\left(\dfrac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^3}\right) \tag{5.18}$$

$$\frac{\partial^2 \log L}{\partial \mu_Y \partial v} = -\sum_{i=1}^{n}\left(\frac{abs(1,-(y_i-\mu_Y)/\sigma_Y)}{\left|y_i-\mu_Y\right|}\right) - \sum_{i=1}^{n}\left(\frac{\sigma_Y\left(\dfrac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^3}\right) \tag{5.19}$$

$$\frac{\partial^2 \log L}{\partial \sigma_X \partial \mu_X} = \left(\begin{array}{l} (v+1)\displaystyle\sum_{i=1}^{n}\left(\left(\frac{x_i-\mu_X}{\sigma_X}\right)\left(\frac{signum(1,-(x_i-\mu_X)/\sigma_X)}{\sigma_X\left|x_i-\mu_X\right|} - \frac{abs(1,-(x_i-\mu_X)/\sigma_X)abs(1,-(x_i-\mu_X))}{\left|x_i-\mu_X\right|^2}\right)\right. \\ \left. -\dfrac{abs(1,-(x_i-\mu_X)/\sigma_X)}{\sigma_X\left|x_i-\mu_X\right|}\right) + \dfrac{2}{\sigma_X}\displaystyle\sum_{i=1}^{n}\left(\frac{x_i-\mu_X}{\sigma_X}\right) - \sum_{i=1}^{n}\left(\frac{\dfrac{2}{\sigma_X}\left(\dfrac{x_i-\mu_X}{\sigma_X}\right)}{\left(\dfrac{y_i-\mu_Y}{\sigma_Y}\right)^2/v}\right) \end{array}\right) \tag{5.20}$$

$$\frac{\partial^2 \log L}{\partial \sigma_X \partial \mu_Y} = v\sum_{i=1}^{n}\left(\frac{\dfrac{2\sigma_Y}{\sigma_X}\left(\dfrac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^3}\right) \tag{5.21}$$

$$\frac{\partial^2 \log L}{\partial \sigma_X \partial \sigma_Y} = \sum_{i=1}^{n}\left(\frac{2\dfrac{\sigma_Y}{\sigma_X}\left(\dfrac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^2/v}\right) \tag{5.22}$$

$$\frac{\partial^2 \log L}{\partial \sigma_X \partial v} = \sum_{i=1}^{n}\left(\left(\frac{x_i-\mu_X}{\sigma_X}\right)\frac{abs(1,-(x_i-\mu_X)/\sigma_X)}{\left|x_i-\mu_X\right|}\right) + \sum_{i=1}^{n}\left(\frac{\dfrac{1}{\sigma_X}\left(\dfrac{x_i-\mu_X}{\sigma_X}\right)^2}{\left(\dfrac{y_i-\mu_Y}{\sigma_Y}\right)^2}\right) \tag{5.23}$$

$$\frac{\partial^2 \log L}{\partial \sigma_Y \partial \mu_X} = \sum_{i=1}^{n}\left(\frac{2\sigma_Y\left(\dfrac{x_i-\mu_X}{\sigma_X}\right)}{(y_i-\mu_Y)^2/v}\right) \tag{5.24}$$

$$\frac{\partial^2 \log L}{\partial\sigma_Y \partial\mu_Y} = -(v+1)\sum_{i=1}^{n}\left(\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)\frac{signum(1,-(y_i-\mu_Y)/\sigma_Y)}{\sigma_Y\left|y_i-\mu_Y\right|}\right) - v\sum_{i=1}^{n}\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{\left(y_i-\mu_Y\right)^3}\right) \tag{5.25}$$

$$\frac{\partial^2 \log L}{\partial\sigma_Y \partial\sigma_X} = \sum_{i=1}^{n}\left(\frac{2\frac{\sigma_Y}{\sigma_X}\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^2/v}\right) \tag{5.26}$$

$$\frac{\partial^2 \log L}{\partial\sigma_Y \partial v} = \sum_{i=1}^{n}\left(\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)\frac{abs(1,-(y_i-\mu_Y)/\sigma_Y)}{\left|y_i-\mu_Y\right|}\right) - \sum_{i=1}^{n}\left(\frac{\sigma_Y\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{\left(y_i-\mu_Y\right)^2}\right) \tag{5.27}$$

$$\frac{\partial \log L}{\partial v \partial\mu_X} = \sum_{i=1}^{n}\log\left|\frac{x_i-\mu_X}{\sigma_X}\right| - \sum_{i=1}^{n}\left(\frac{abs(1,-(x_i-\mu_X)/\sigma_X)}{\left|x_i-\mu_X\right|}\right) + \sum_{i=1}^{n}\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^2}\right) \tag{5.28}$$

$$\frac{\partial \log L}{\partial v \partial\mu_Y} = -\sum_{i=1}^{n}\left(\frac{abs(1,-(y_i-\mu_Y)/\sigma_Y)}{\left|y_i-\mu_Y\right|}\right) - \sigma_Y^2\sum_{i=1}^{n}\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^3}\right) \tag{5.29}$$

$$\frac{\partial \log L}{\partial v \partial\sigma_X} = \sum_{i=1}^{n}\left(\left(\frac{x_i-\mu_X}{\sigma_X}\right)\frac{abs(1,-(x_i-\mu_X)/\sigma_X)}{\left|x_i-\mu_X\right|}\right) + \frac{1}{\sigma_X}\sum_{i=1}^{n}\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^2}\right) \tag{5.30}$$

and

$$\frac{\partial \log L}{\partial v \partial \sigma_Y} = -\sum_{i=1}^{n}\left(\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)\frac{abs(1,-(y_i-\mu_Y)/\sigma_Y)}{\left|y_i-\mu_Y\right|}\right) - \sigma_Y \sum_{i=1}^{n}\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^2}\right) \quad (5.31)$$

From (5.1) to (5.31), one can exhibit the elements of the Fisher Information matrix are

$$E\left(-\frac{\partial^2 \log L}{\partial \mu_X^2}\right) = -\frac{v}{\sigma_X}\sum_{i=1}^{n}\left(E\left(\frac{signum(1,-(x_i-\mu_X)/\sigma_X)}{\left|x_i-\mu_X\right|}\right) - E\left(\frac{abs(1,-(x_i-\mu_X)/\sigma_X)^2}{\left|x_i-\mu_X\right|^2}\right)\right)$$
$$+\frac{n}{\sigma_X} - \frac{n}{\sigma_X}\sum_{i=1}^{n}\left(E\left(\frac{1}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^2 / v}\right)\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial \mu_Y^2}\right) = \frac{v+1}{\sigma_Y}\sum_{i=1}^{n}\left(E\left(\frac{signum(1,-(y_i-\mu_Y)/\sigma_Y)}{\left|y_i-\mu_Y\right|}\right) - E\left(\frac{abs(1,-(y_i-\mu_Y)/\sigma_Y)^2}{\left|y_i-\mu_Y\right|}\right)\right)$$
$$+3v\sum_{i=1}^{n}\left(E\left(\frac{\sigma_Y\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^4}\right)\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial \sigma_X^2}\right) = \left(\begin{array}{l} -\frac{n}{2\sigma_X^2} - (v+1)\sum_{i=1}^{n}\left(E\left(\left(\frac{\left|x_i-\mu_X\right|}{\sigma_X^3}\right) signum(1,-(x_i-\mu_X)/\sigma_X)\right)\right. \\ \left. -E\left(\left(\frac{x_i-\mu_X}{\sigma_X^2}\right)\frac{abs(1,-(x_i-\mu_X)/\sigma_X)}{\left|x_i-\mu_X\right|}\right)\right) \\ -\frac{4}{\sigma_X^3}\sum_{i=1}^{n} E\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2 + \frac{4}{\sigma_X^3}\sum_{i=1}^{n} E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^2 / v}\right) \end{array}\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial \sigma_Y^2}\right)=\left(\begin{array}{l} -\frac{n}{2\sigma_Y^2}-(v+1)\left(\sum_{i=1}^{n}\left(E\left(\left(\frac{\left|y_i-\mu_Y\right|}{\sigma_Y^3}\right)\frac{signum(1,-(y_i-\mu_Y)/\sigma_Y)}{\left|y_i-\mu_Y\right|}\right)\right.\right. \\ \left.\left.+E\left(\left(\frac{y_i-\mu_Y}{\sigma_Y^3}\right)\frac{abs(1,-(y_i-\mu_Y)/\sigma_Y)}{\left|y_i-\mu_Y\right|}\right)\right)\right)+\sum_{i=1}^{n}E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{\left(y_i-\mu_Y\right)^2/v}\right)\end{array}\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial v^2}\right)=-\frac{n}{v}+\frac{n}{2}\Psi(1,v/2)$$

$$E\left(-\frac{\partial^2 \log L}{\partial \mu_X \partial \mu_Y}\right)=-\frac{1}{\sigma_Y}\sum_{i=1}^{n}E\left(\frac{2\left(\frac{x_i-\mu_X}{\sigma_X}\right)}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^3/v}\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial \mu_X \partial \sigma_X}\right)=-v\left(\sum_{i=1}^{n}\left(E\left(\left(\frac{x_i-\mu_X}{\sigma_X}\right)\frac{signum\left(1,-\left(x_i-\mu_X\right)/\sigma_X\right)}{\sigma_X\left|x_i-\mu_X\right|}\right)\right)\right)+$$

$$\sum_{i=1}^{n}E\left(\frac{x_i-\mu_X}{\sigma_X^2}\right)+\sum_{i=1}^{n}E\left(\frac{\frac{x_i-\mu_X}{\sigma_X^2}}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^2/v}\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial \mu_X \partial \sigma_Y}\right)=-\frac{2}{\sigma_Y}\sum_{i=1}^{n}E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^3/v}\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial \mu_Y \partial \mu_X}\right) = -\frac{2}{\sigma_Y}\sum_{i=1}^{n} E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^3 / v}\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial \mu_Y \partial \sigma_X}\right) = -\frac{2v\sigma_Y}{\sigma_X}\sum_{i=1}^{n} E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^3}\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial \mu_Y \partial \sigma_Y}\right) = (v+1)\sum_{i=1}^{n}\left(E\left(\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)\frac{signum(1,-(y_i-\mu_Y)/\sigma_Y)}{\sigma_Y\left|y_i-\mu_Y\right|}\right)\right) + v\sum_{i=1}^{n}\left(E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^3}\right)\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial \sigma_X \partial \mu_X}\right) = \left(\begin{array}{l} -(v+1)\sum_{i=1}^{n}\left(E\left(\left(\frac{x_i-\mu_X}{\sigma_X}\right)\left(\frac{signum(1,-(x_i-\mu_X)/\sigma_X)}{\sigma_X\left|x_i-\mu_X\right|} - \frac{abs(1,-(x_i-\mu_X)/\sigma_X)abs(1,-(x_i-\mu_X))}{\left|x_i-\mu_X\right|^2}\right)\right)\right) \\ -E\left(\frac{abs(1,-(x_i-\mu_X)/\sigma_X)}{\sigma_X\left|x_i-\mu_X\right|}\right)\Bigg)\Bigg) + \frac{2}{\sigma_X}\sum_{i=1}^{n}E\left(\frac{x_i-\mu_X}{\sigma_X}\right) - \frac{2}{\sigma_X}\sum_{i=1}^{n}E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^2 / v}\right) \end{array}\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial \sigma_X \partial \mu_Y}\right) = -\frac{2v\sigma_Y}{\sigma_X}\sum_{i=1}^{n} E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^3}\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial \sigma_X \partial \sigma_Y}\right) = -2\frac{\sigma_Y}{\sigma_X}\sum_{i=1}^{n} E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^2 / v}\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial\sigma_Y \partial\mu_X}\right) = -2\sigma_Y \sum_{i=1}^{n} E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)}{(y_i-\mu_Y)^2/v}\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial\sigma_Y \partial\mu_Y}\right) = (v+1)\sum_{i=1}^{n}\left(E\left(\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)\frac{signum(1,-(y_i-\mu_Y)/\sigma_Y)}{\sigma_Y\left|y_i-\mu_Y\right|}\right)\right) + v\sum_{i=1}^{n} E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^3}\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial\sigma_Y \partial\sigma_X}\right) = -2\frac{\sigma_Y}{\sigma_X}\sum_{i=1}^{n} E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^2/v}\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial\sigma_Y \partial v}\right) = \sum_{i=1}^{n}\left(E\left(\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)\frac{abs(1,-(y_i-\mu_Y)/\sigma_Y)}{\left|y_i-\mu_Y\right|}\right)\right) - \sigma_Y\sum_{i=1}^{n} E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^2}\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial v \partial\mu_X}\right) = -\sum_{i=1}^{n} E\left(\log\left|\frac{x_i-\mu_X}{\sigma_X}\right|\right) + \sum_{i=1}^{n} E\left(\frac{abs(1,-(x_i-\mu_X)/\sigma_X)}{\left|x_i-\mu_X\right|}\right) - \sum_{i=1}^{n} E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^2}\right)$$

$$E\left(-\frac{\partial \log L}{\partial v \partial\mu_Y}\right) = \sum_{i=1}^{n} E\left(\frac{abs(1,-(y_i-\mu_Y)/\sigma_Y)}{\left|y_i-\mu_Y\right|}\right) + \sigma_Y^2\sum_{i=1}^{n} E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^3}\right)$$

$$E\left(-\frac{\partial^2 \log L}{\partial v \partial\sigma_X}\right) = -\sum_{i=1}^{n} E\left(\left(\frac{x_i-\mu_X}{\sigma_X}\right)\frac{abs(1,-(x_i-\mu_X)/\sigma_X)}{\left|x_i-\mu_X\right|}\right) - \frac{1}{\sigma_X}\sum_{i=1}^{n} E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^2}\right)$$

and

$$E\left(-\frac{\partial^2 \log L}{\partial v \partial \sigma_Y}\right)=\sum_{i=1}^{n} E\left(\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)\frac{abs(1,-(y_i-\mu_Y)/\sigma_Y)}{\left|y_i-\mu_Y\right|}\right)+\sigma_Y\sum_{i=1}^{n} E\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{(y_i-\mu_Y)^2}\right)$$

Result 5.2: As a second approach, the authors realized that the computational complexity of the maximum likelihood estimation of the proposed distribution due to the high non-linearity of parameters. Hence, Non-linear programming technique is adopted by using Constrained Maximum likelihood method to estimate the parameters of the Bivariate mixture of t-normal distribution and the idea is to Maximizing the log-likelihood function from (5.1) under some restriction and parameter constraints and it given as follows:

$$\log L(x,y;\mu_X,\mu_Y,\sigma_X,\sigma_Y,v)=$$
$$\left\{\begin{array}{l} -\frac{n}{2}\log(2\pi)+\frac{nv}{2}\log(v/2)-n\log(\Gamma(v/2))-n\log(\sigma_X\sigma_Y)+v\sum_{i=1}^{n}\log\left|\frac{x_i-\mu_X}{\sigma_X}\right| \\ -(v+1)\sum_{i=1}^{n}\log\left|\frac{y_i-\mu_Y}{\sigma_Y}\right|-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^2/v}\right) \end{array}\right\} \tag{5.32}$$

Subject to the constraints

$$\log L(x,y;\mu_X,\mu_Y,\sigma_X,\sigma_Y,v)=$$
$$\left\{\begin{array}{l} -\frac{n}{2}\log(2\pi)+\frac{nv}{2}\log(v/2)-n\log(\Gamma(v/2))-n\log(\sigma_X\sigma_Y)+v\sum_{i=1}^{n}\log\left|\frac{x_i-\mu_X}{\sigma_X}\right| \\ -(v+1)\sum_{i=1}^{n}\log\left|\frac{y_i-\mu_Y}{\sigma_Y}\right|-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{\left(\frac{x_i-\mu_X}{\sigma_X}\right)^2}{\left(\frac{y_i-\mu_Y}{\sigma_Y}\right)^2/v}\right) \end{array}\right\}\leq 0$$

$$-\infty<\mu_X,\mu_Y<+\infty$$

$$\sigma_X, \sigma_Y, v > 0$$

Conclusion

This paper introduces a novel Bivariate mixture of the student's *t* and Gaussian normal distributions, referred to as the *t*-Normal distribution. The distribution's characteristics are examined in detail, revealing two key properties. First, the conditional variances within the *t*-Normal distribution exhibit heteroscedasticity. Second, although the correlation between the standard normal and the student's *t* variables is zero—indicating that they are uncorrelated—the variables are not independent. This suggests that the population (normal) and sampling (Student's *t*) variables are not associated in a linear sense, yet they retain a dependency structure.

Future work will explore the distributions of the sum, difference, and ratio of these two variables. Additionally, the multivariate extension of the *t*-Normal mixture will be considered in subsequent research.

References

[1] S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 6th ed. San Diego, CA: Academic Press (2000).

[2] P. Prudnikov, Y. A. Brychkov, and O. I. Marichev, *Integrals and Series*, Vols. 1, 2, and Amsterdam: Gordon and Breach Science Publishers (1986).

[3] S. Stigler, *Seven Pillars of Statistical Wisdom*. Cambridge, MA: Harvard University Press (2016).

[4] S. Stigler, *The History of Statistics*. Cambridge, MA: Harvard University Press (1986).

[5] Abraham de Moivre, *The Doctrine of Chances*, 2nd ed. London, UK (1738).

[6] R. A. Fisher, "Applications of 'Student's' distribution," *Metron*, vol. 5, pp. 90–104 (1925).

[7] W. S. Gosset, "Letter No. 2 to R. A. Fisher," Guinness Son & Co Ltd, Dublin (1962). [Private Collection].

[8] E. S. Pearson, "'Student' as statistician," *Biometrika*, vol. 30, pp. 210–250 (1939).

[9] W. S. Gosset, "The probable error of a mean," *Biometrika*, vol. 4, pp. 1–24, 1908; "Tables for estimating the probability that the mean of a unique sample of observations lies between -∞ and any given distance of the mean of the population from which the sample is drawn," (1917).

[10] L. McMullen, "'Student' as a man," *Biometrika*, vol. 30, pp. 204–210 (1939).

[11] Y. Sayyari, "New refinements of Shannon's entropy upper bounds," *J. Inf. Optim. Sci.*, vol. 42, no. 8, pp. 1869–1883 (2021).

Received April, 2025