On the limiting distributions of the Johnson System of Frequency curves - A Geometric Analysis

"Presentation Short Course: Beyond Beta and Applications"

***November 20th, 2018**, La Sapienza*

J. René van Dorp*

** Department of Engineering Management and Systems Engineering,*
School of Engineering and Applied Science,The George Washington University,
800 22nd Street, N.W., Suite 2800,
Washington D.C. 20052. E-mail: dorpjr@gwu.edu
http://www.seas.gwu.edu/~dorpjr

OUTLINE

1. **INTRODUCTION**
2. GENERAL THEORY
3. THE S_B-SYSTEM
4. THE S_L-SYSTEM
5. THE S_U-SYSTEM
6. SUMMARY AND CONCLUSIONS
7. SELECTED REFERENCES

- **N.L. Johnson (1917 - 2004)** is one of the leaders of the 20th century Statistics (specializing in **Statistical Distribution Theory and Applications**).
- Johnson studied at the University College London (UCL) at the tail end of **Karl Pearson's (the father of the beta distribution** among other important distributions) life in the 1930's.
- Johnson's joint contributions with other contributors to the field of Statistics are plentiful, **publishing over 180 papers, monographs and tables.**
- In particular his contributions with **Samuel Kotz (1930-2010)** are prolific, publishing 17 books, including the ***Distribution in Statistics*** volumes. Jointly they served as editor of the 10-volume ***Encyclopedia of Statistics*** and the three-volume ***Breakthrough in Statistics***.


Norman L. Johnson


Samuel Kotz

- In his thesis, **summarized in Johnson (1949)**, Johnson developed ***a pioneering* system of transformations of the standard normal distribution** which received substantial popularity in the second half of the 20th century.

- **Johnson's System of Frequency curves have been extensively investigated in the statistical literature for some 50 years or more.** In his own words, he considered his discovery of these systems of frequency curves the **"main piece of work"** on his own – see, Read (2004).

- Of note is that Johnson's systems of frequency curves **cover all of the twelve Pearson's curves** – see, e.g., Patil *et al.* (1984) – and were developed in the late 1940's before the introduction of (even the most primitive) computers into statistical practice.

- **Computations related to the Johnson's systems of frequency curves are quite ingenuous and involved**, and they were originally carried out with old-fashioned graphical calculators, **which required long hours and even days of patient calculating** of something that nowadays may take less than 1 second.

- As such, **the Johnson systems of frequency curves** have **paved the way for introduction of computation intensive methodology** and the use of non-standard functions (such as the hyperbolic and Jacobi functions) **in statistical practice.**

- Herein, **a geometric rediscovery** of the Johnson's systems of frequency curves is provided

- Through this geometric rediscovery light is shedding **on the limiting distributions** of the three families/systems of frequency curves proposed by Johnson (1949), **a topic which received little to no attention.**

- In the process, **the geometric interpretation of Johnson's four parameters** is explained as well as **their influence on the shape of the frequency curves**, which was one of **Johnson's primary interests as it relates to these translations systems.**

OUTLINE

1. INTRODUCTION
2. **GENERAL THEORY**
3. THE S_B-SYSTEM
4. THE S_L-SYSTEM
5. THE S_U-SYSTEM
6. SUMMARY AND CONCLUSIONS
7. SELECTED REFERENCES

- **Gaussian random variable** $Z \sim N(0,1)$ with probability density function (pdf)

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, z \in (-\infty, \infty)$$

- Consider **the transformation functions $f(\cdot)$ and $f^{-1}(\cdot)$** with ***translation* parameters $\gamma, \xi \in \mathbb{R}$, $\delta, \lambda > 0$**

$$Z = \gamma + \delta f\Big(\frac{X-\xi}{\lambda}\Big) \Leftrightarrow X = \xi + \lambda f^{-1}\Big(\frac{Z-\gamma}{\delta}\Big).$$

- The function $f(\cdot)$ translates ***X*** **into** ***Z*** utilizes and $f^{-1}(\cdot)$ translates ***Z*** **into** ***X***. Thus the function $f(\cdot)$ (or $f^{-1}(\cdot)$) needs to be **a monotonic function.**

- With the condition that $f(\cdot)$ is a **non-decreasing and differentiable**, the pdf $p(x)$ of the random variate X follows as

$$p(x) = \delta \times f'\{(x-\xi)/\lambda\} \times \phi[\gamma + \delta f\{(x-\xi)/\lambda\}].$$

Note the importance of **the gradient/derivative of the transformation function $f(\cdot)$** in this equation.

- Johnson (1949) proposes the following **three types of translation systems:**

1) **The S_B - bounded support system:**

$$f(\cdot):(0,1)\rightarrow(-\infty,\infty),\text{ where } f(y)=\ln\{y/(1-y)\}$$

2) **The S_L - lognormal system:**

$$f(\cdot):(0,\infty)\rightarrow(-\infty,\infty),\text{ where } f(y)=\ln(y)$$

3) **The S_U - unbounded support system:**

$$f(\cdot):(-\infty,\infty)\rightarrow(-\infty,\infty),\text{ where } f(y)=\ln(y+\sqrt{y^2+1}).$$

- **The ranges of the functions $f(\cdot)$** share **the support of the Gaussian pdf $\phi(z)$.**

- **The domains of the functions $f(\cdot)$** share **the support of the pdf $p(x)$.**

- The symmetry of the functions $f(\cdot)$ in the S_B-System and the S_U-System combined with the symmetry of the Gaussian pdf implies that both the S_B and S_U systems allow for symmetric frequency curves.

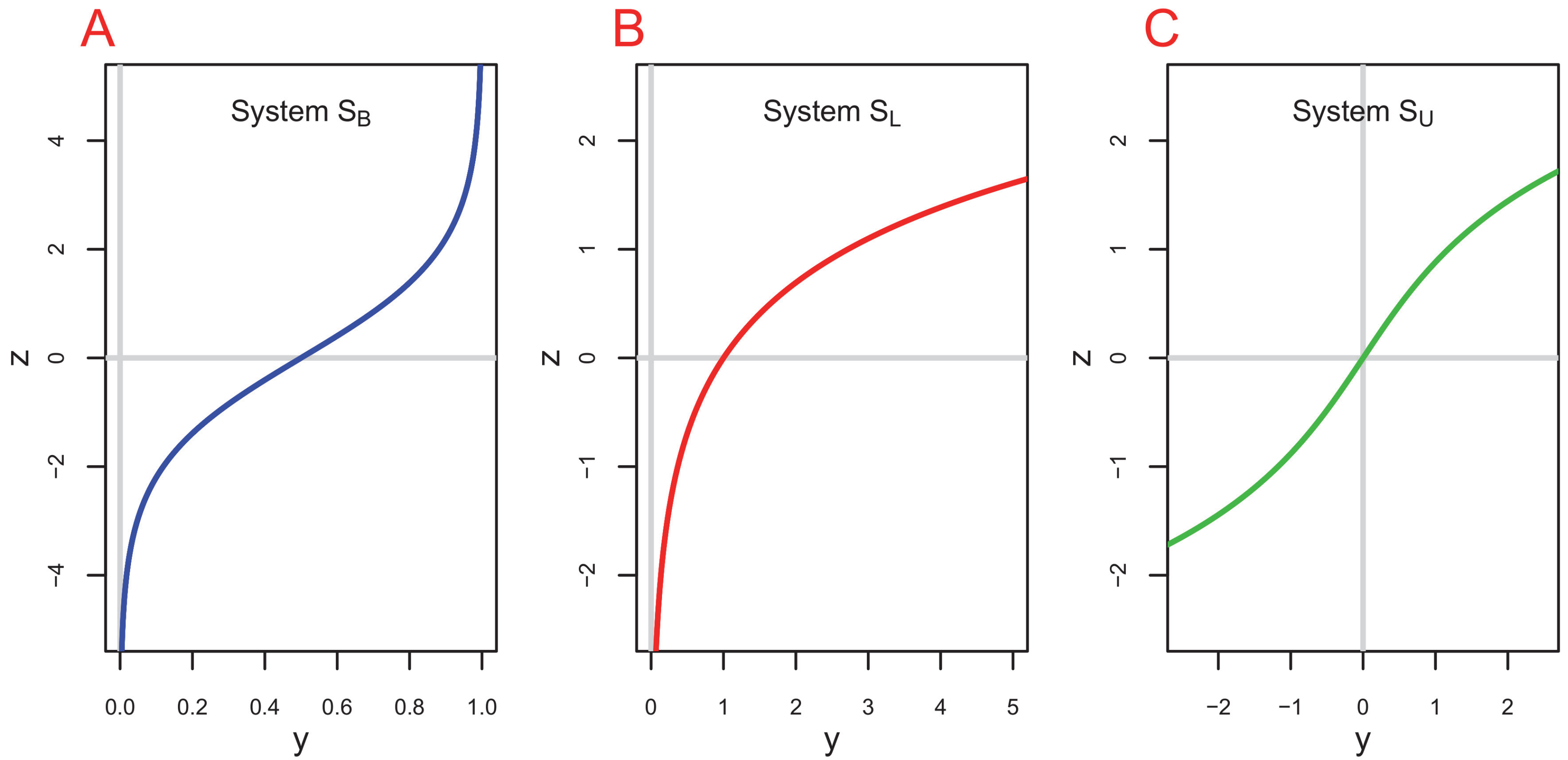
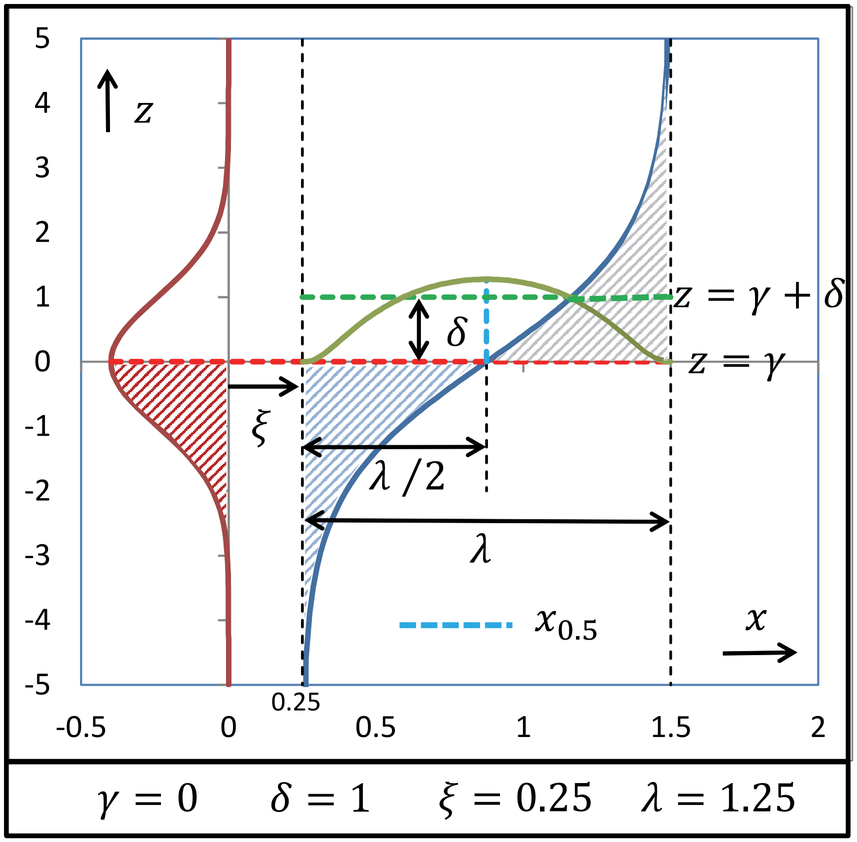


Figure. Johnson's curves $f(\cdot)$: A: S_B-System; B: S_L-System; C: S_U-System.

- **The variety of shapes of pdf's** given **Johnson's functions $f(\cdot)$ combined with the Gaussian pdf $\phi(z)$** is as large as that of **the whole Pearson system** (see, e.g., Stuart and Ord (1994)).

OUTLINE

1. INTRODUCTION
2. GENERAL THEORY
3. **THE S_B-SYSTEM**
4. THE S_L-SYSTEM
5. THE S_U-SYSTEM
6. SUMMARY AND CONCLUSIONS
7. SELECTED REFERENCES

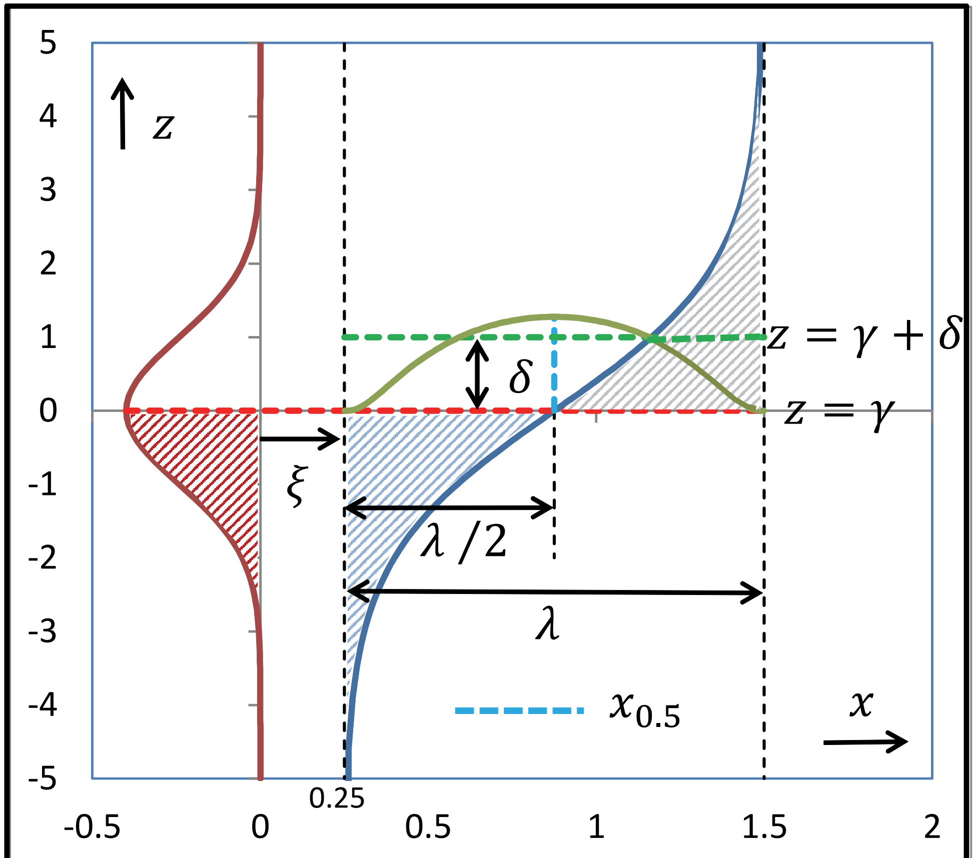


$f(\cdot):(0,1)\rightarrow(-\infty,\infty)$, where
$f(y)= \ln\{y/(1-y)\}$

The asymptotes of **the translation function $\gamma+\delta f\{(x-\xi)/\lambda\}$** at $x=\xi$ and $x=\xi+\lambda$ define the support of the **S_B-pdf's $p(x)$**.

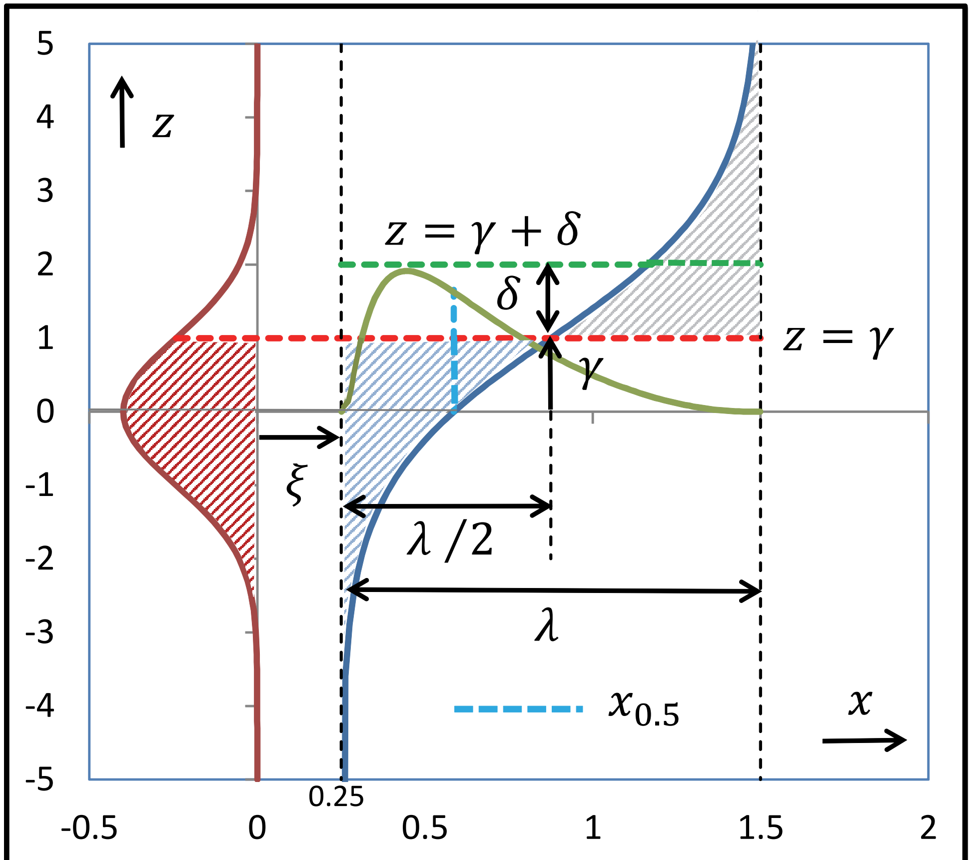
The **symmetric** blue and the gray shaded areas both capture 50% of the probability mass of Z which translates to a **median** at the support midpoint $x_{0.5}=\xi+\lambda/2$.

$$\gamma+\delta f\{(x_{0.5}-\xi)/\lambda\}=0 \Rightarrow x_{0.5}=\lambda y_{0.5}+\xi, \text{ where } y_{0.5}=\frac{exp(-\frac{\gamma}{\delta})}{1+exp(-\frac{\gamma}{\delta})}.$$



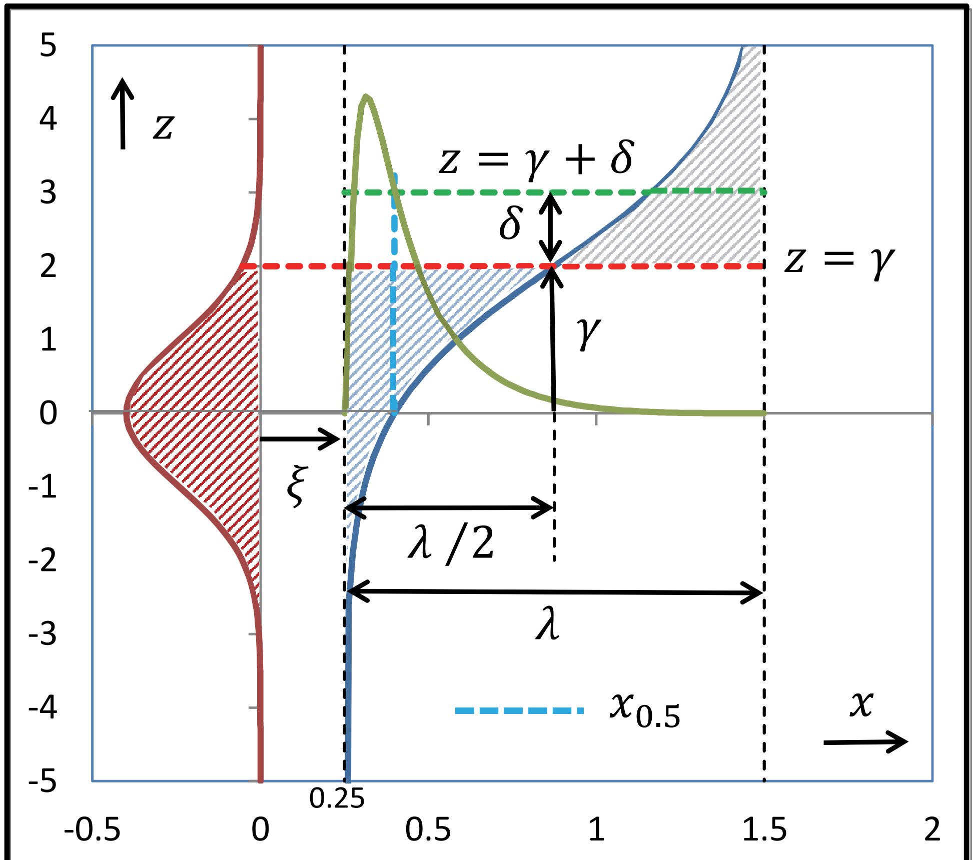
$\gamma = 0 \quad \delta = 1 \quad \xi = 0.25 \quad \lambda = 1.25$

- By increasing γ but keeping δ fixed, **the gradients $f'(\cdot)$ of the translation function remain the same**, but probability **captured by blue area** increases.



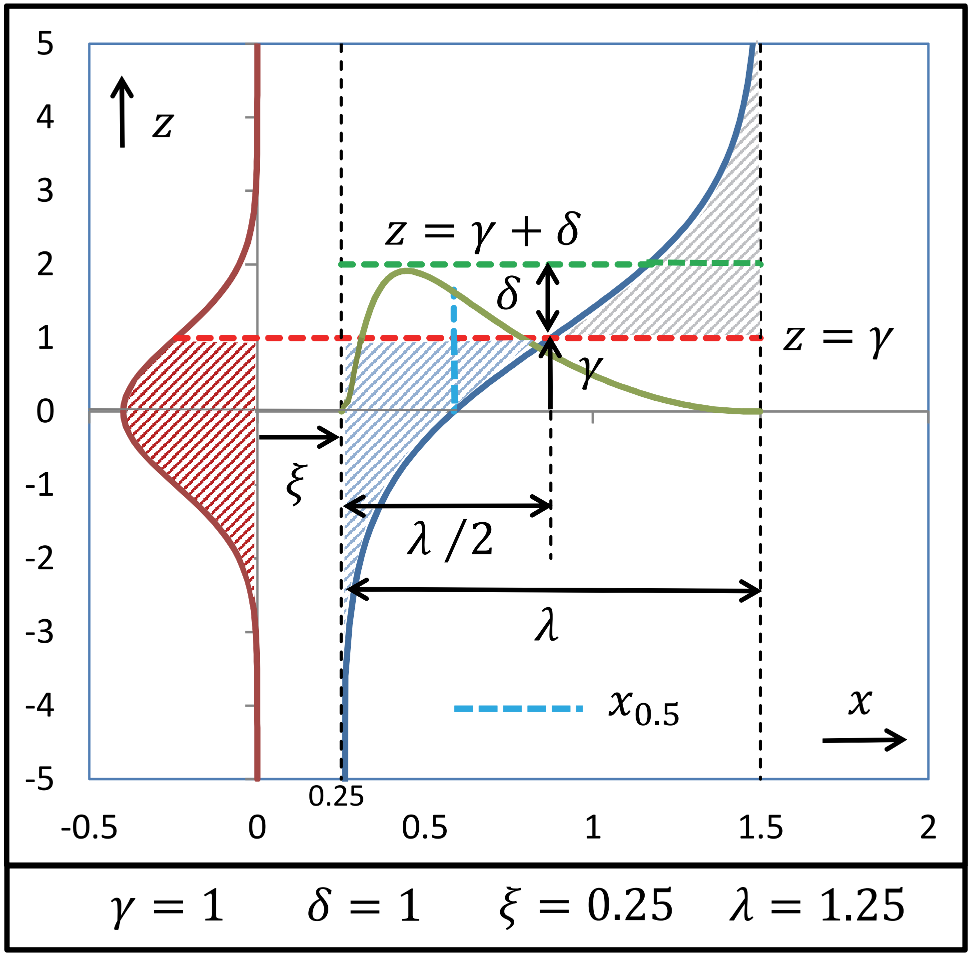
$\gamma = 1$ $\delta = 1$ $\xi = 0.25$ $\lambda = 1.25$

- By increasing γ but keeping δ fixed, **the gradients $f'(\cdot)$ of the translation function remain the same**, but probability **captured by blue area** increases.

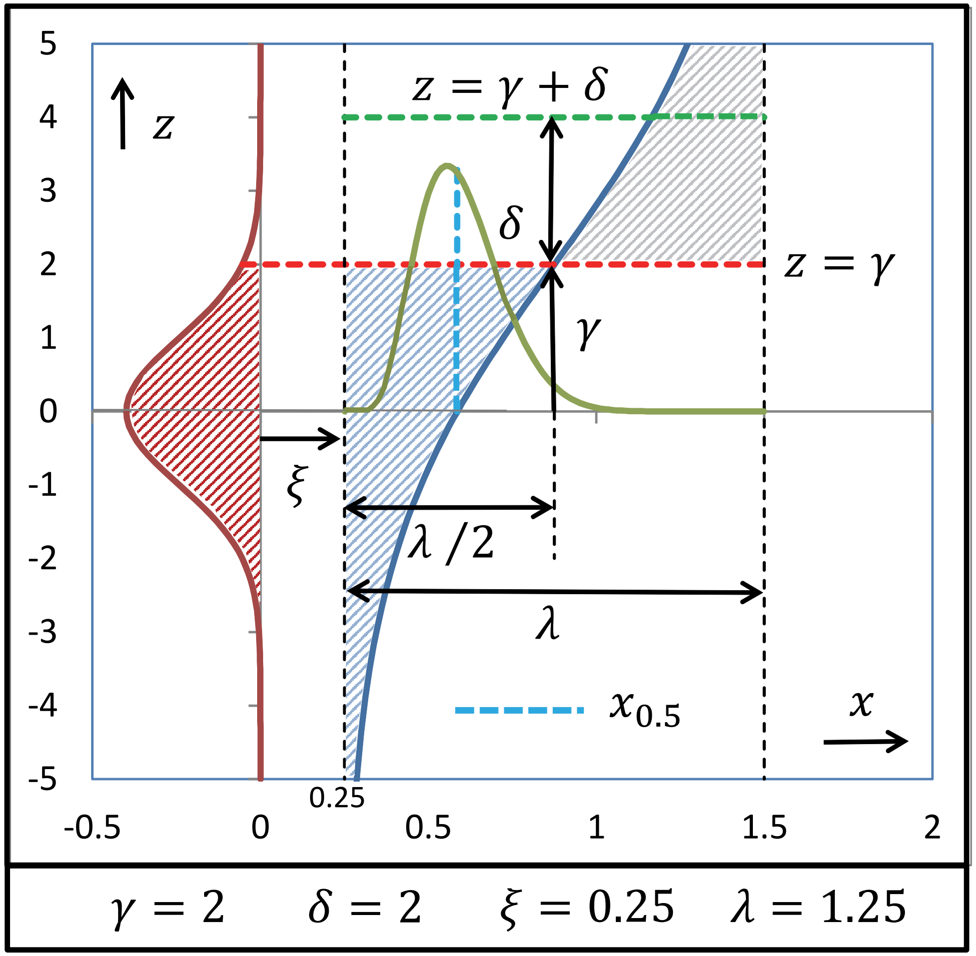


$\gamma = 2 \quad \delta = 1 \quad \xi = 0.25 \quad \lambda = 1.25$

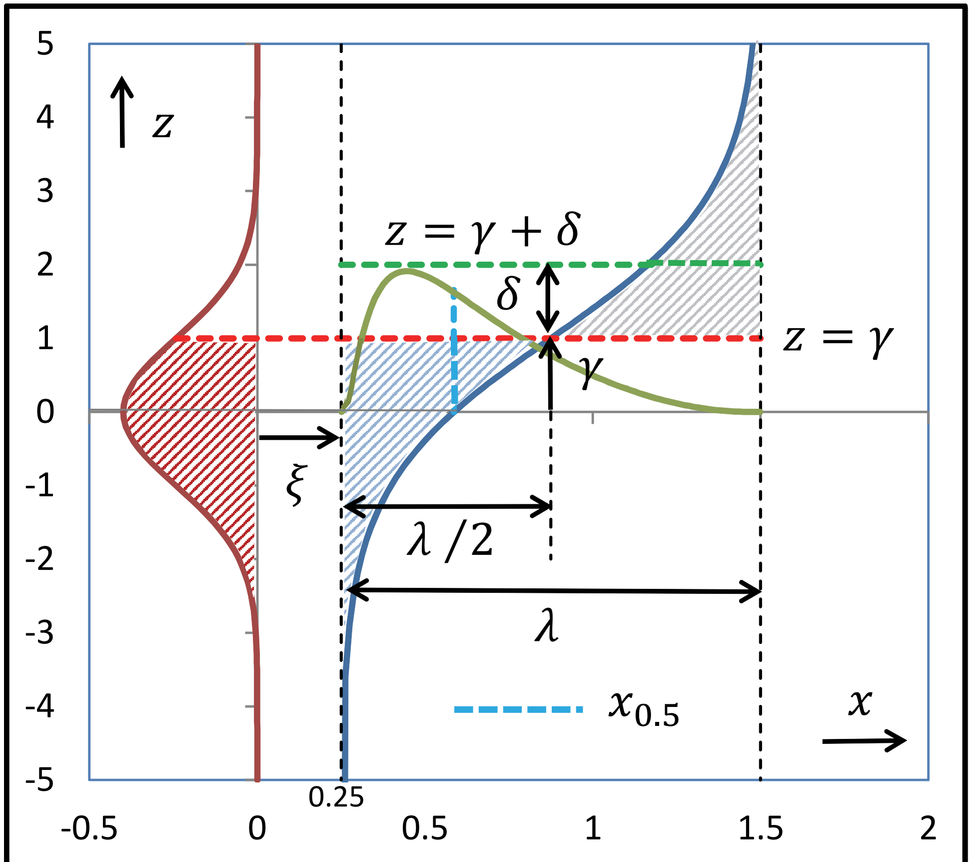
- S_B pdf $p(x) \rightarrow$ **point mass 1 at ξ** as $\gamma \rightarrow \infty$, **keeping δ fixed.**
  S_B pdf $p(x) \rightarrow$ **point mass 1 at $\xi + \lambda$** as $\gamma \rightarrow -\infty$, **keeping δ fixed.**



- Letting $\delta \to \infty$, but **keeping γ/δ fixed**, the median $x_{0.5}$ remains while the tangent line of $\gamma + \delta f\{(x-\xi)/\lambda\}$ converges to **vertical line at $x = x_{0.5}$.**

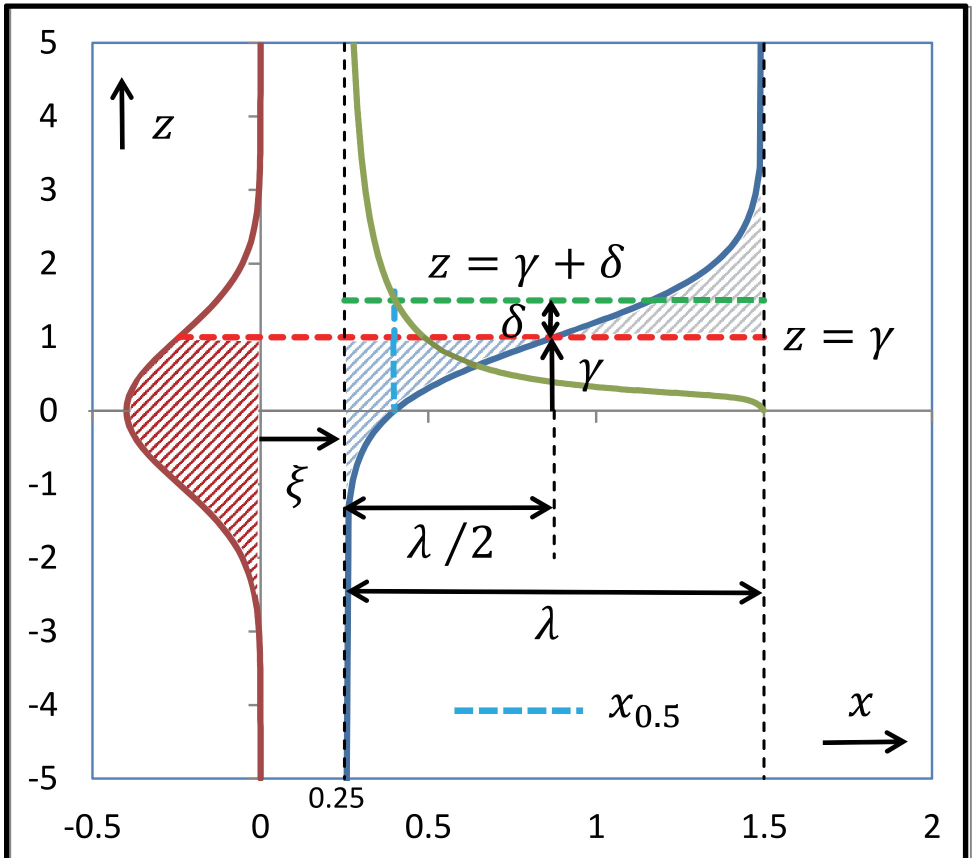


- S_B Pdf $p(x) \rightarrow$ **single point mass at $x_{0.5}$** as $\delta \rightarrow \infty$, **keeping γ/δ fixed.**



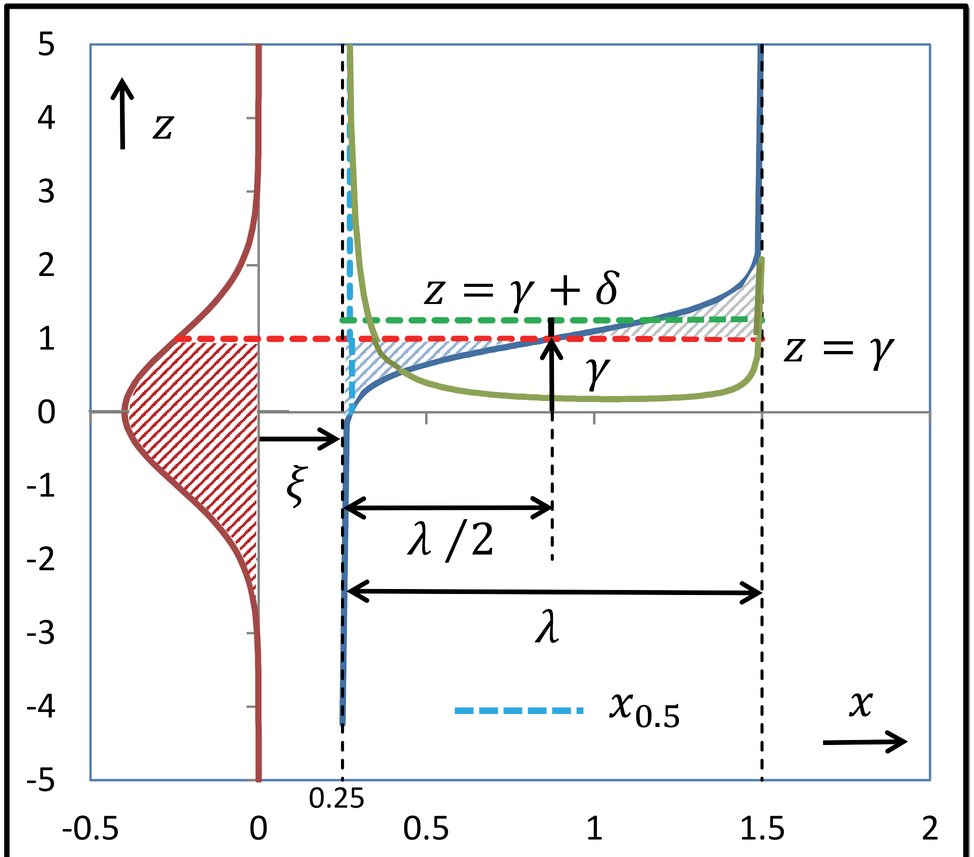
$\gamma = 1 \quad \delta = 1 \quad \xi = 0.25 \quad \lambda = 1.25$

- Letting $\delta \downarrow 0$, but **keeping γ fixed**, the tangent line of $\gamma + \delta f\{(x-\xi)/\lambda\}$ at $(\xi + \lambda/2, \gamma)$ converges to **the horizontal line $z = \gamma$.**



$\gamma = 1 \quad \delta = 0.5 \quad \xi = 0.25 \quad \lambda = 1.25$

- Letting $\delta \downarrow 0$, but **keeping γ fixed**, the tangent line of $\gamma + \delta f\{(x-\xi)/\lambda\}$ at $(\xi + \lambda/2, \gamma)$ converges to **the horizontal line $z = \gamma$.**



$\gamma = 1 \quad \delta = 0.25 \quad \xi = 0.25 \quad \lambda = 1.25$

- S_B Pdf $p(x) \rightarrow$ **Point mass $\Phi(\gamma)$ at ξ** and **point mass $1 - \Phi(\gamma)$ at $\xi + \lambda$** as $\delta \downarrow 0$, **keeping γ fixed.** Thus it converges to a **Bernoulli RV.**

- The Johnson S_B-pdfs $p(x)$ share **the same limiting distributions** as those of **the classical beta pdf** (Karl Pearson (1895))

$$g(x|\alpha,\beta,\xi,\lambda)=\frac{1}{\lambda}(x-\xi)^{\alpha-1}(\xi+\lambda-x)^{\beta-1}, x\in(\xi,\xi+\lambda),$$

the more recently discovered **Kumaraswamy (1980) distribution**

$$g(x|\alpha,\beta,\xi,\lambda)=\frac{1}{\lambda}\alpha\beta(x-\xi)^{\alpha-1}(\xi+\lambda-x^{\alpha})^{\beta-1}, x\in(\xi,\xi+\lambda),$$

and **Two-Sided Power distribution** by Kotz and van Dorp (2002) with pdf:

$$g(x|\theta,n,\xi,\lambda)=\frac{n}{\lambda}\times\begin{cases}(\frac{x-\xi}{\theta-\xi})^{n-1}, & \xi<x<\xi+\theta,\\ (\frac{\xi+\lambda-x}{\xi+\lambda-\theta})^{n-1}, & \theta<x<\xi+\lambda.\end{cases}$$

- **The latter two distributions** have **closed form cdf's** and the former two do not, which is the distinct advantage of these latter two discoveries.

- **The 30-year time-lag** between these discoveries is a testament to **the pioneering work** of both **N.L Johnson and Karl Pearson**.

OUTLINE

1. INTRODUCTION
2. GENERAL THEORY
3. THE S_B-SYSTEM
4. **THE S_L-SYSTEM**
5. THE S_U-SYSTEM
6. SUMMARY AND CONCLUSIONS
7. SELECTED REFERENCES

- Setting $\delta = \sigma^{-1}, \gamma = -\mu/\sigma,\ \xi = 0$ and $\lambda = 1$ in

$$Z = \gamma + \delta f\left(\frac{X-\xi}{\lambda}\right) \Leftrightarrow X = \xi + \lambda f^{-1}\left(\frac{Z-\gamma}{\delta}\right)$$

with $f(y) = \ln(y)$ it follows that:

$$Z = -\frac{\mu}{\sigma} + \frac{1}{\sigma}\ln(X) \Leftrightarrow \ln(X) = \sigma Z + \mu \sim N(\mu, \sigma).$$

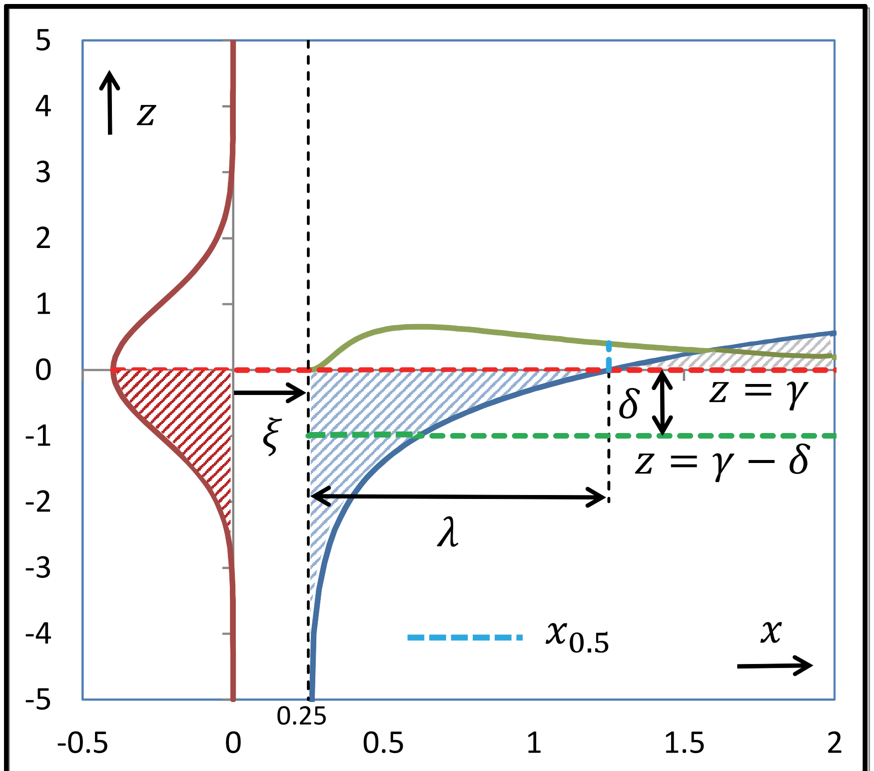
In other words, X is **lognormally distributed** (see. Gaddum (1945))

- For $\xi \in \mathbb{R}$ and $\lambda > 0$, it follows with $\delta = \sigma^{-1}$ and $\gamma = -\mu/\sigma$ that

$$\ln(X - \xi) = \sigma Z + \mu + \ln(\lambda) \sim N\{\mu + \ln(\lambda), \sigma\}. \tag{12}$$

Thus, **without loss of generality**, Johnson (1949) sets $\lambda = 1$ and the S_L-System reduces to **a three-parameter family of distributions** parameterized by $\xi, \gamma \in \mathbb{R}$ and $\delta > 0$.

- **Behavior of $p(x)$ as a function of $\gamma \in \mathbb{R}, \delta > 0$** is **quite different** then as a function of **the traditional lognormal $\mu \in \mathbb{R}, \sigma > 0$ parameters.**



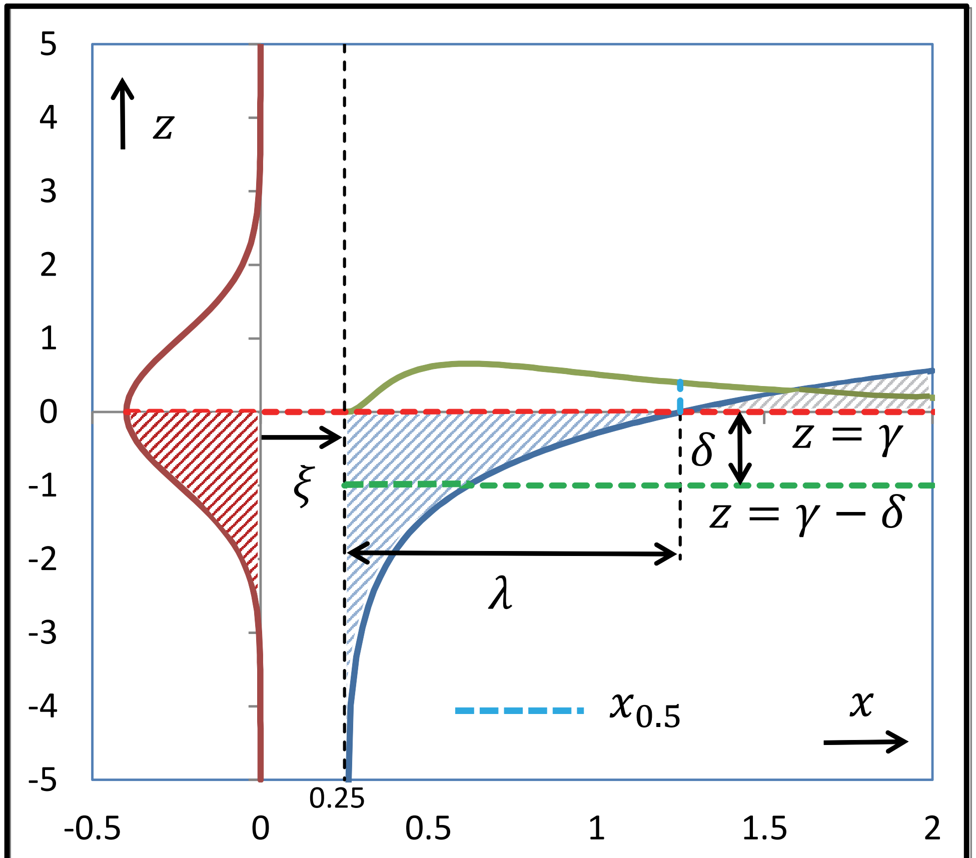
$\gamma = 0$ $\delta = 1$ $\xi = 0.25$ $\lambda = 1$

$f(\cdot) : (0, \infty) \rightarrow (-\infty, \infty)$,
where $f(y) = \ln(y)$

The asymptote of **the translation function $\gamma + \delta f\{(x - \xi)/\lambda\}$** as a function of the variable x at $\boldsymbol{x = \xi}$ defines the support $[\xi, \infty)$ of **the S_L-pdf's $p(x)$.**

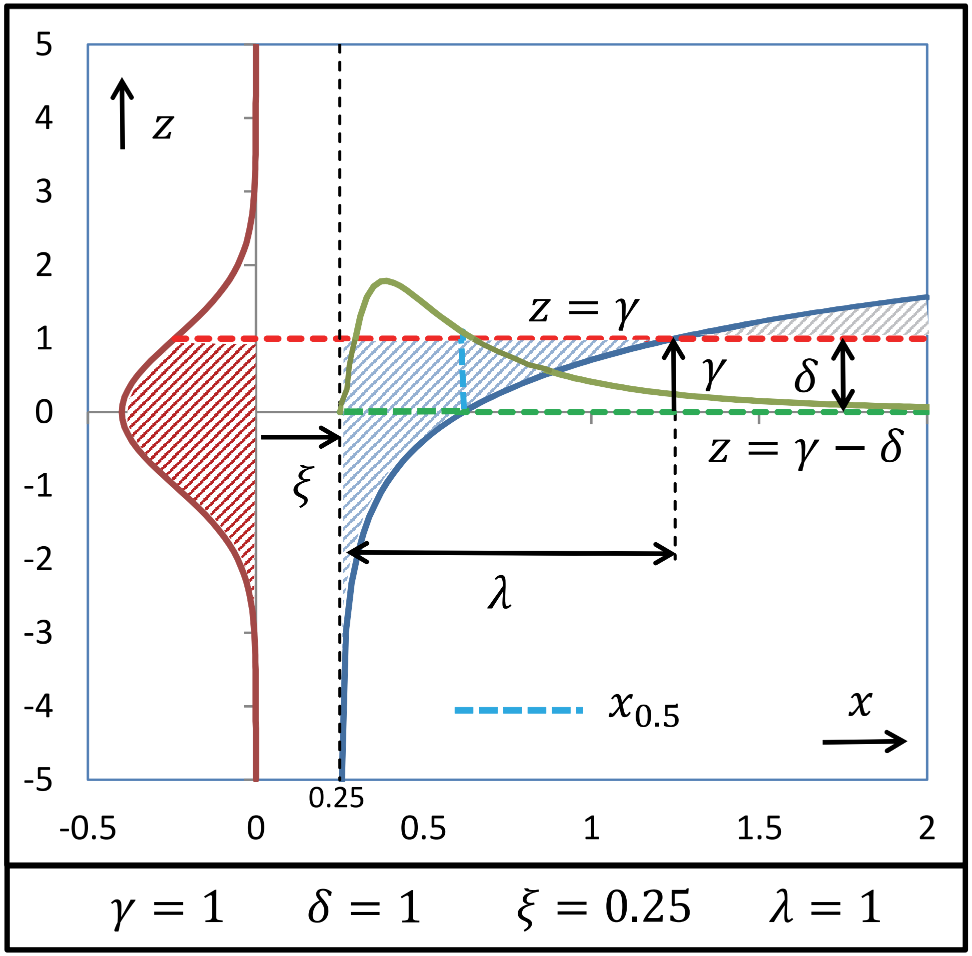
The blue and the gray shaded both areas capture 50% of the probability mass of Z which translates to a median at $\boldsymbol{x_{0.5} = \xi + \lambda}$.

$$\gamma + \delta f\{(x_{0.5} - \xi)/\lambda\} = 0 \Rightarrow x_{0.5} = \lambda y_{0.5} + \xi, \text{ where } y_{0.5} = exp(-\gamma/\delta)$$

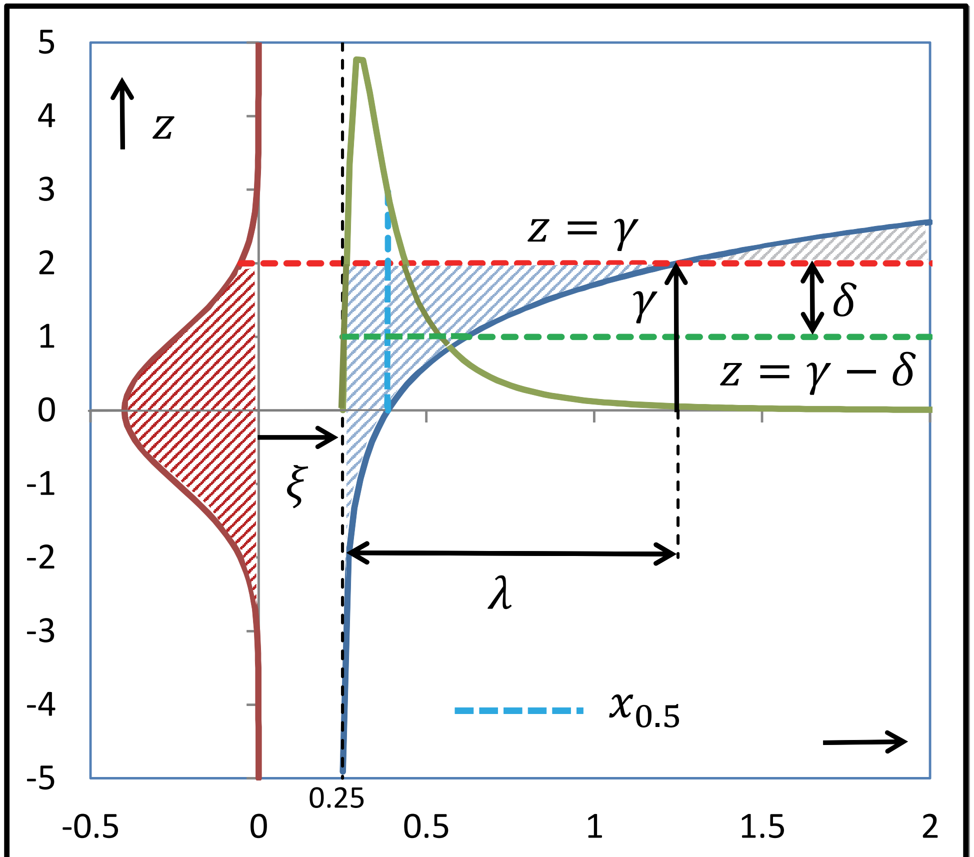


$\gamma = 0 \quad \delta = 1 \quad \xi = 0.25 \quad \lambda = 1$

- By increasing γ but keeping δ fixed, **the gradients $f'(\cdot)$ of the translation function remain the same**, but probability **captured by blue area** increases.

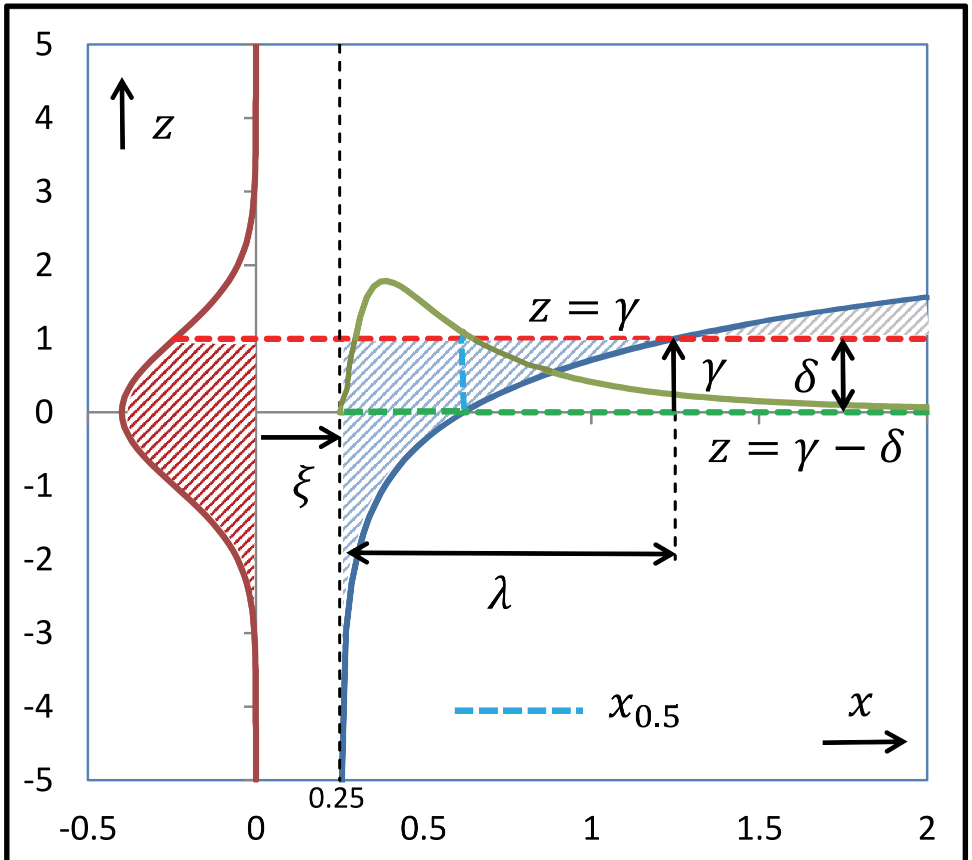


- By increasing γ but keeping δ fixed, **the gradients $f'(\cdot)$ of the translation function remain the same**, but probability **captured by blue area** increases.



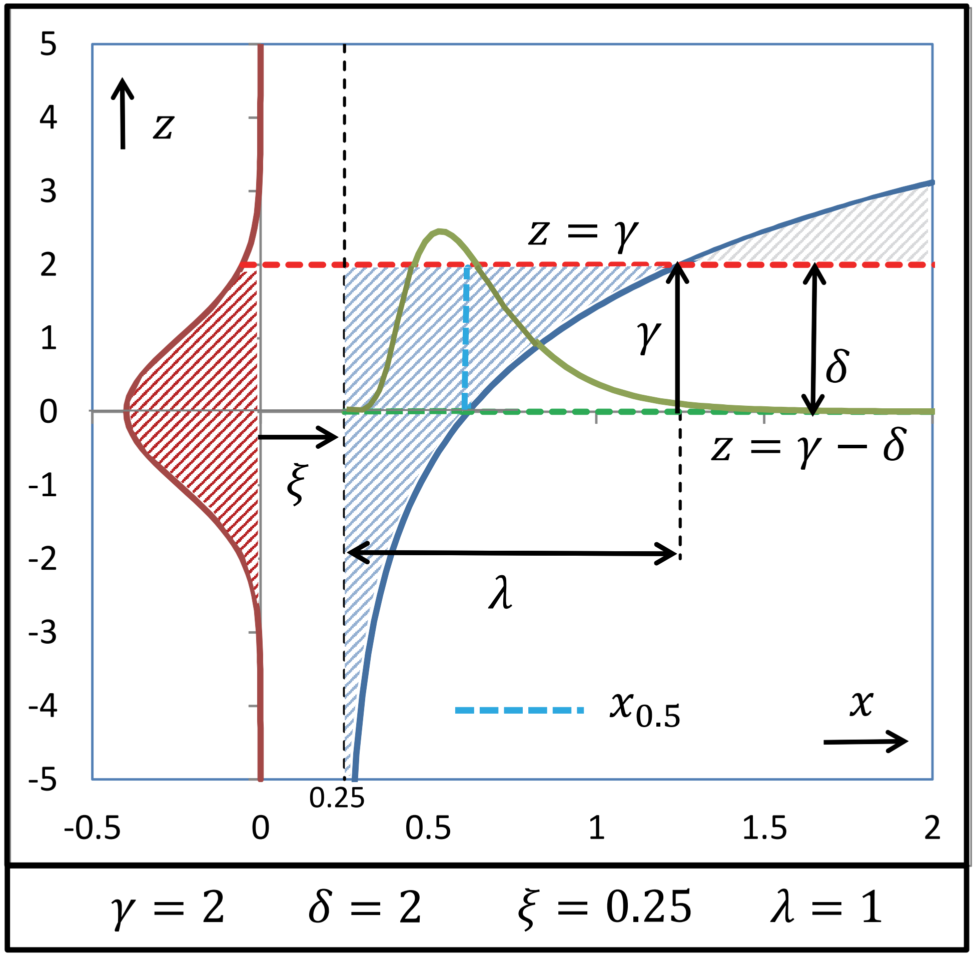
$\gamma = 2 \quad \delta = 1 \quad \xi = 0.25 \quad \lambda = 1$

- S_L Pdf $p(x) \to$ single point mass at ξ as $\gamma \to \infty$, **keeping δ fixed.**
  S_L Pdf $p(x) \to$ probability mass of 1 at ∞ as $\gamma \to -\infty$, **keeping δ fixed.**

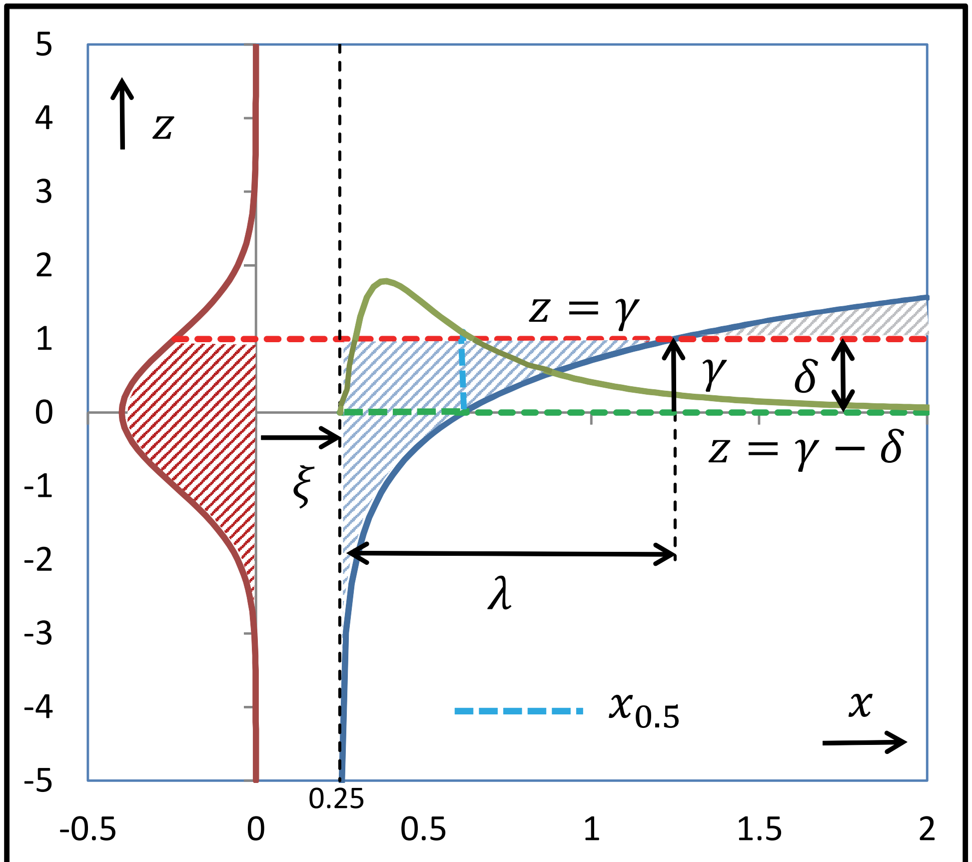


$\gamma = 1$	$\delta = 1$	$\xi = 0.25$	$\lambda = 1$

- Letting $\delta \to \infty$, but **keeping γ/δ fixed**, the median $x_{0.5}$ remains while the tangent line of $\gamma + \delta f\{(x-\xi)/\lambda\}$ converges to **vertical line at $x = x_{0.5}$**.

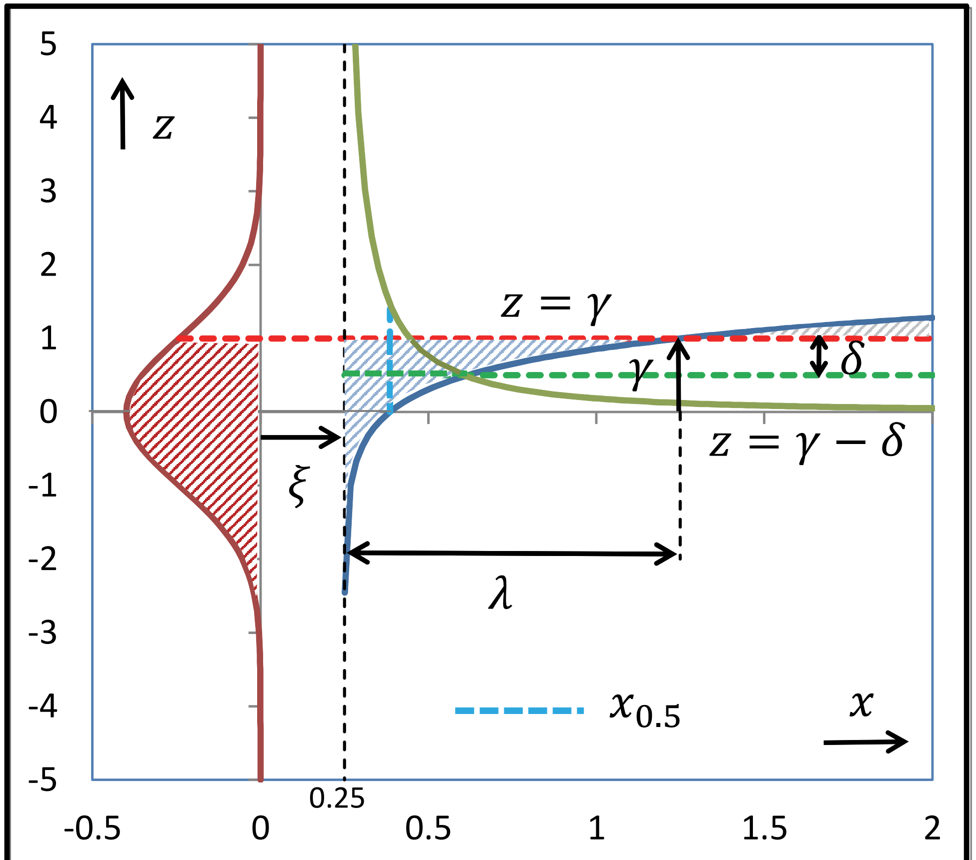


- S_L Pdf $p(x) \to$ **single point mass at $x_{0.5}$** as $\delta \to \infty$, **keeping γ/δ fixed.**



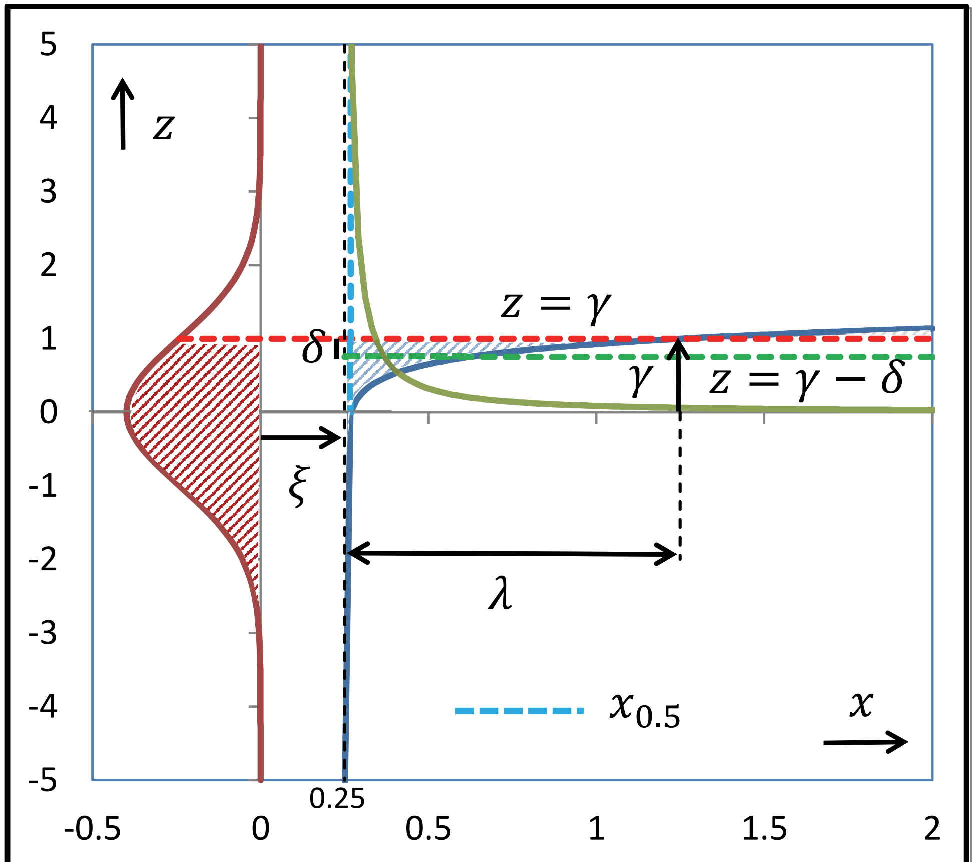
$\gamma = 1$ $\delta = 1$ $\xi = 0.25$ $\lambda = 1$

- Letting $\delta \downarrow 0$, but **keeping γ fixed**, the tangent line of $\gamma + \delta f\{(x-\xi)/\lambda\}$ at $(\xi + \lambda, \gamma)$ converges to **the horizontal line $z = \gamma$.**



$\gamma = 1 \quad \delta = 0.5 \quad \xi = 0.25 \quad \lambda = 1$

- Letting $\delta \downarrow 0$, but **keeping γ fixed**, the tangent line of $\gamma + \delta f\{(x-\xi)/\lambda\}$ at $(\xi + \lambda, \gamma)$ converges to **the horizontal line $z = \gamma$.**

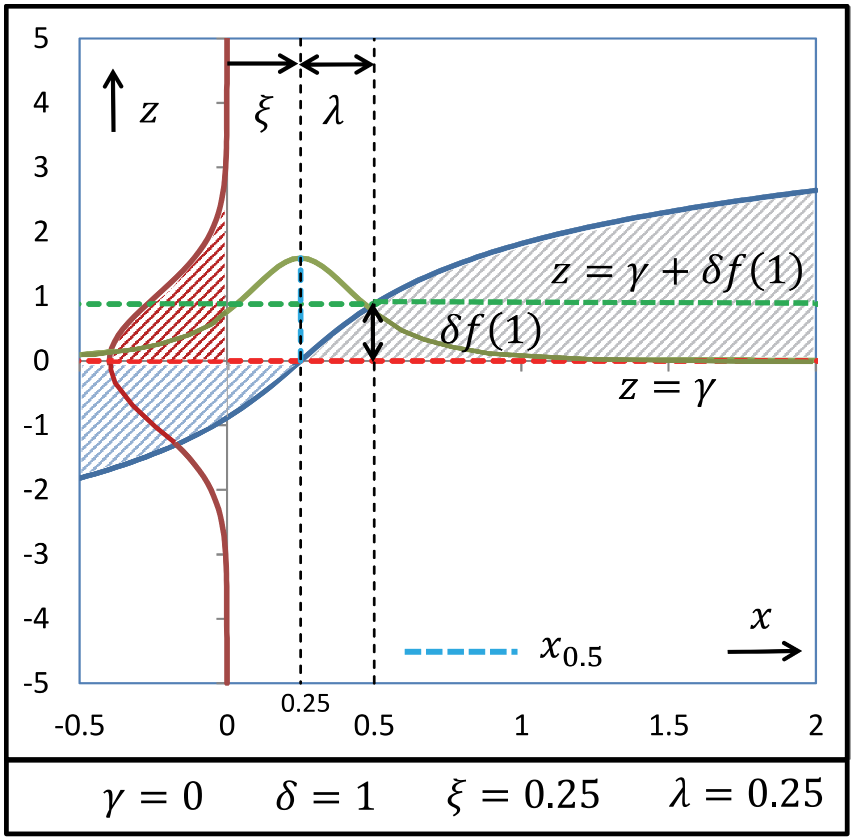


$\gamma = 1 \quad \delta = 0.25 \quad \xi = 0.25 \quad \lambda = 1$

- S_L Pdf $p(x) \rightarrow$ **Point mass $\Phi(\gamma)$ at ξ** and **probability mass $1 - \Phi(\gamma)$ at ∞** as $\delta \downarrow 0$, **keeping γ fixed.**

OUTLINE

1. INTRODUCTION
2. GENERAL THEORY
3. THE S_B-SYSTEM
4. THE S_L-SYSTEM
5. **THE S_U-SYSTEM**
6. SUMMARY AND CONCLUSIONS
7. SELECTED REFERENCES

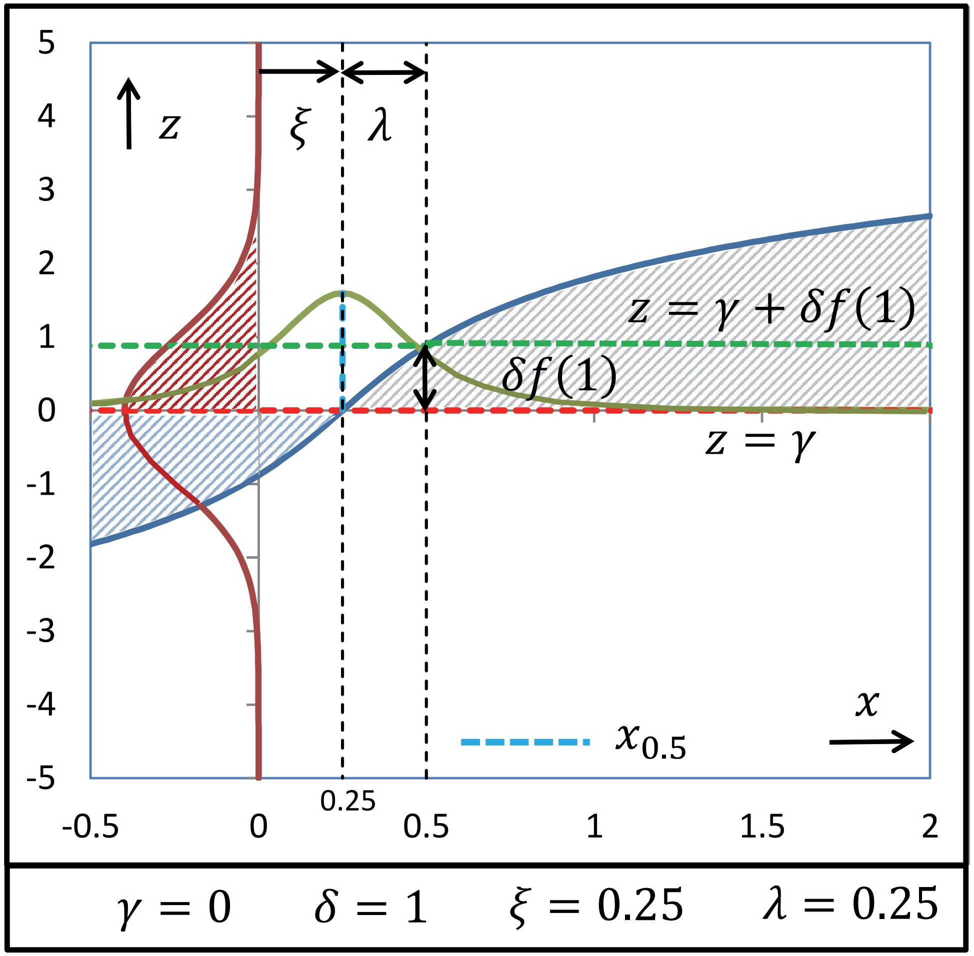


$f(\cdot):(-\infty,\infty)\to(-\infty,\infty)$,
where $f(y)=\ln(y+\sqrt{y^2+1})$

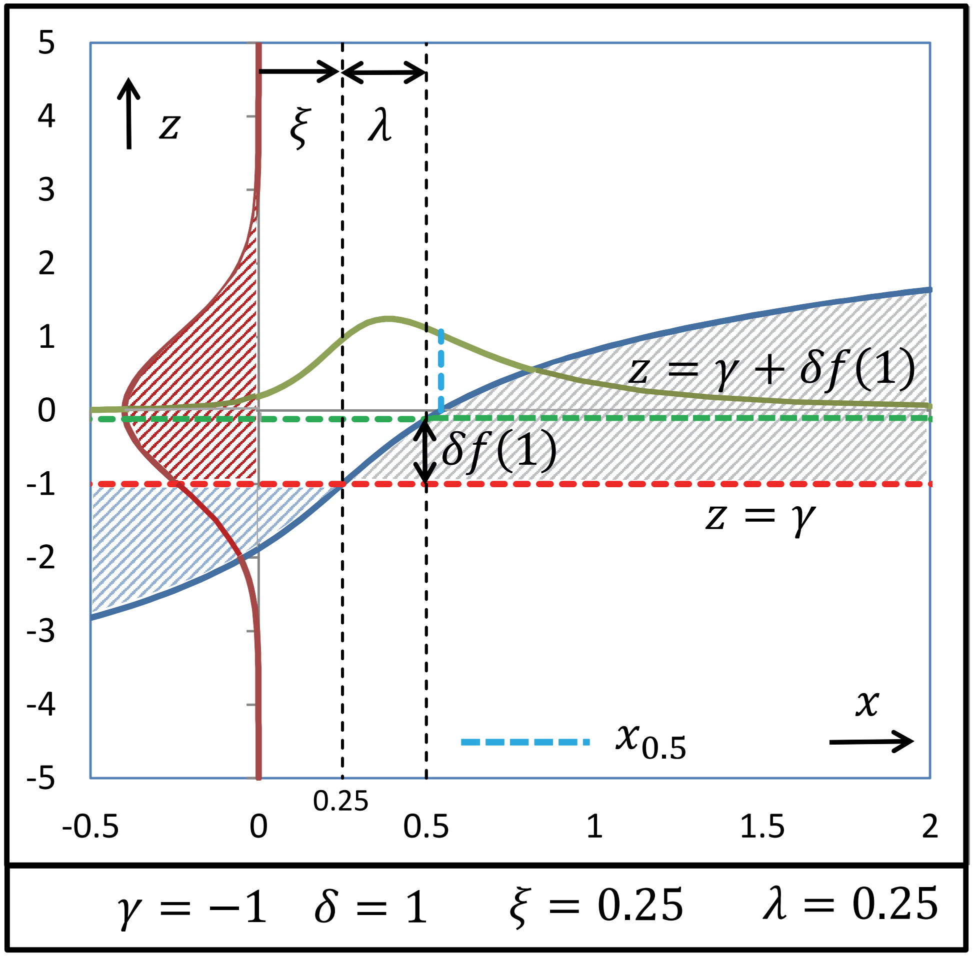
The translation function
$\boldsymbol{\gamma+\delta f\{(x-\xi)/\lambda\}}$ as a function of the variable x defines the support $(-\infty,\infty)$ of **the S_U-pdf's $p(x)$.**

The **symmetric** blue and the gray shaded areas both capture 50% of the probability mass of Z which translates to a median at $\boldsymbol{x_{0.5}=\xi}$.

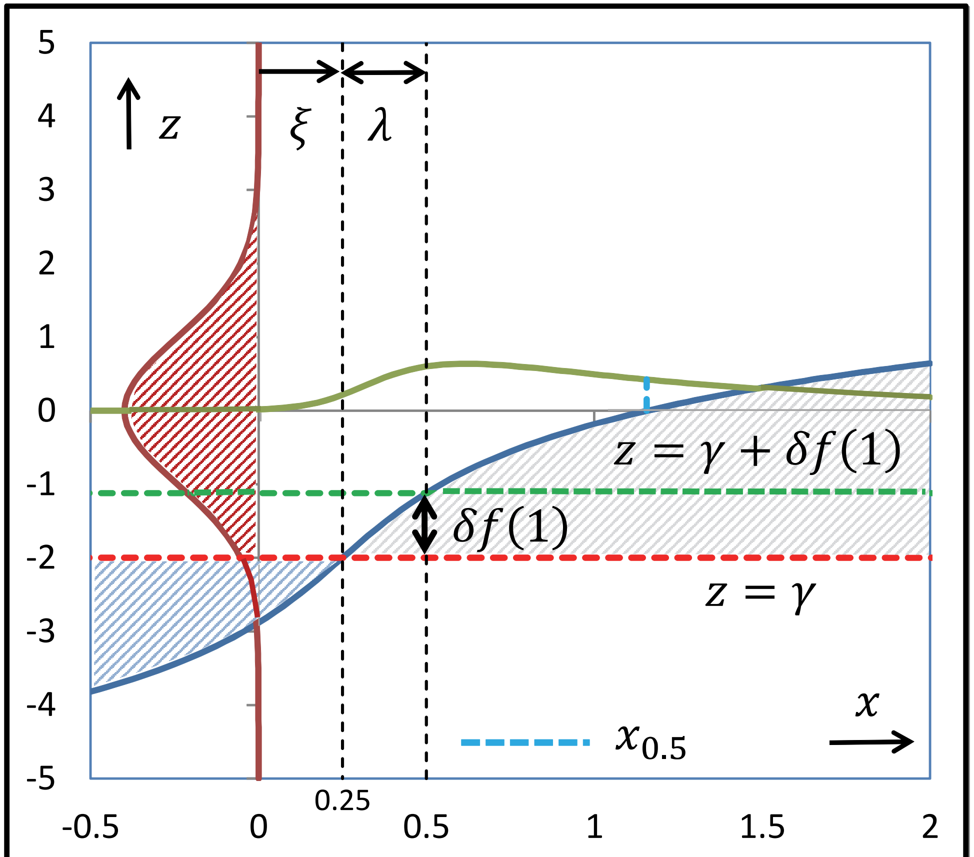
$$\gamma+\delta f\{(x_{0.5}-\xi)/\lambda\}=0 \Rightarrow x_{0.5}=\lambda y_{0.5}+\xi,\ y_{0.5}=\frac{\{exp(-\gamma/\delta)\}^2-1}{2exp(-\gamma/\delta)}$$



- By decreasing γ but keeping δ fixed, **the gradients $f'(\cdot)$ of the translation function remains the same**, but probability **captured by gray area** increases.



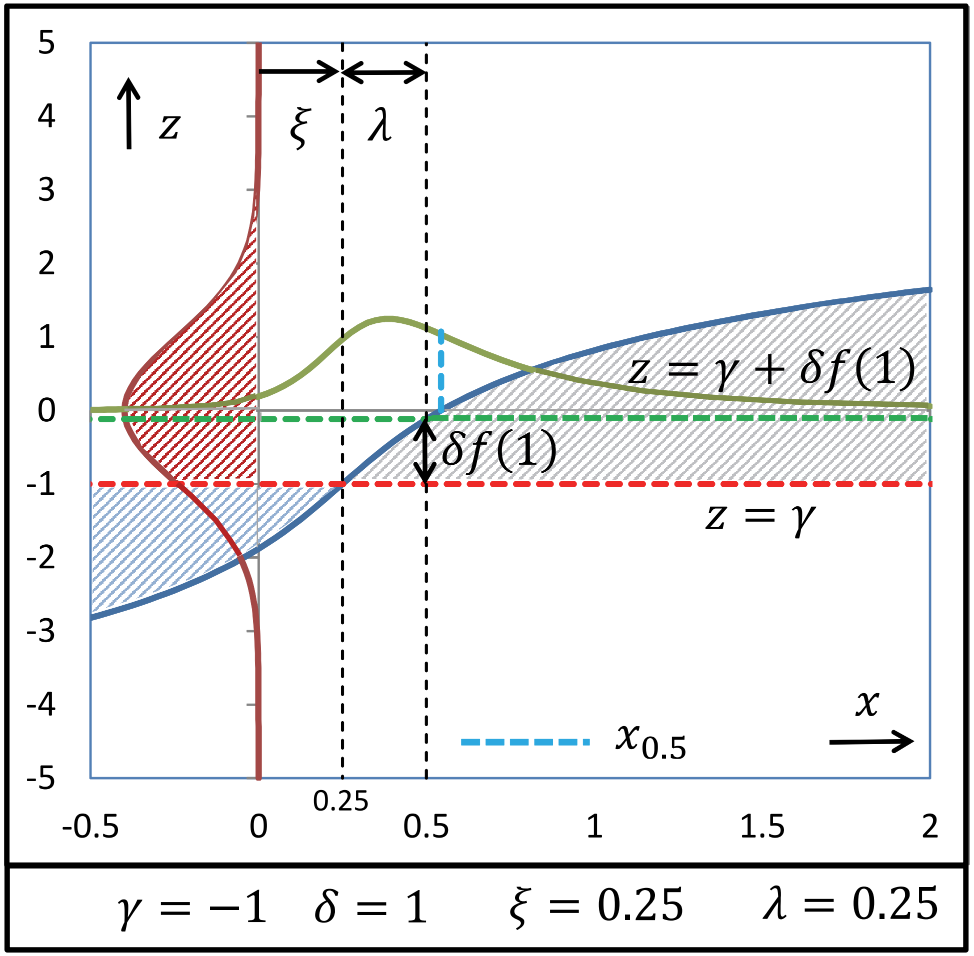
- By decreasing γ but keeping δ fixed, **the gradients $f'(\cdot)$ of the translation function remains the same**, but probability **captured by gray area** increases.



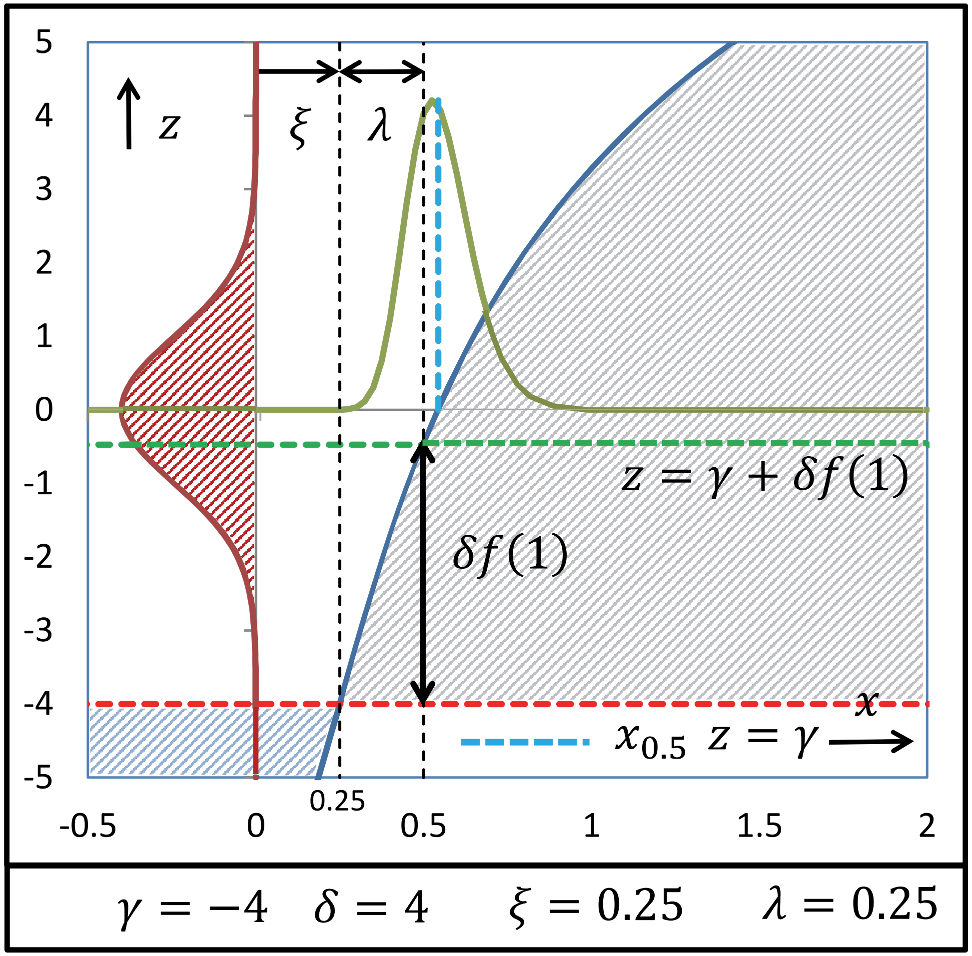
$\gamma = -2 \quad \delta = 1 \quad \xi = 0.25 \quad \lambda = 0.25$

- S_U Pdf $p(x) \rightarrow$ **prob. mass of 1 at ∞** as $\gamma \rightarrow -\infty$, **keeping δ fixed.**

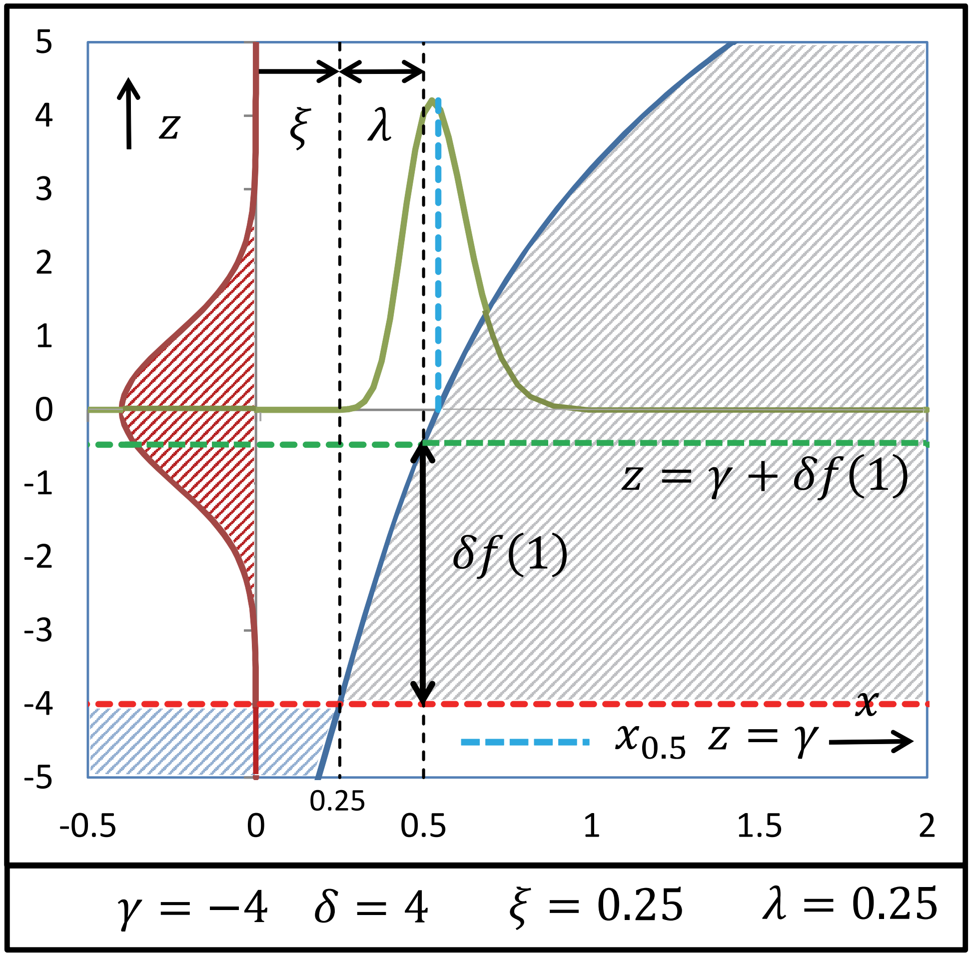
  S_U Pdf $p(x) \rightarrow$ **prob. mass of 1 at $-\infty$** as $\gamma \rightarrow \infty$, **keeping δ fixed.**



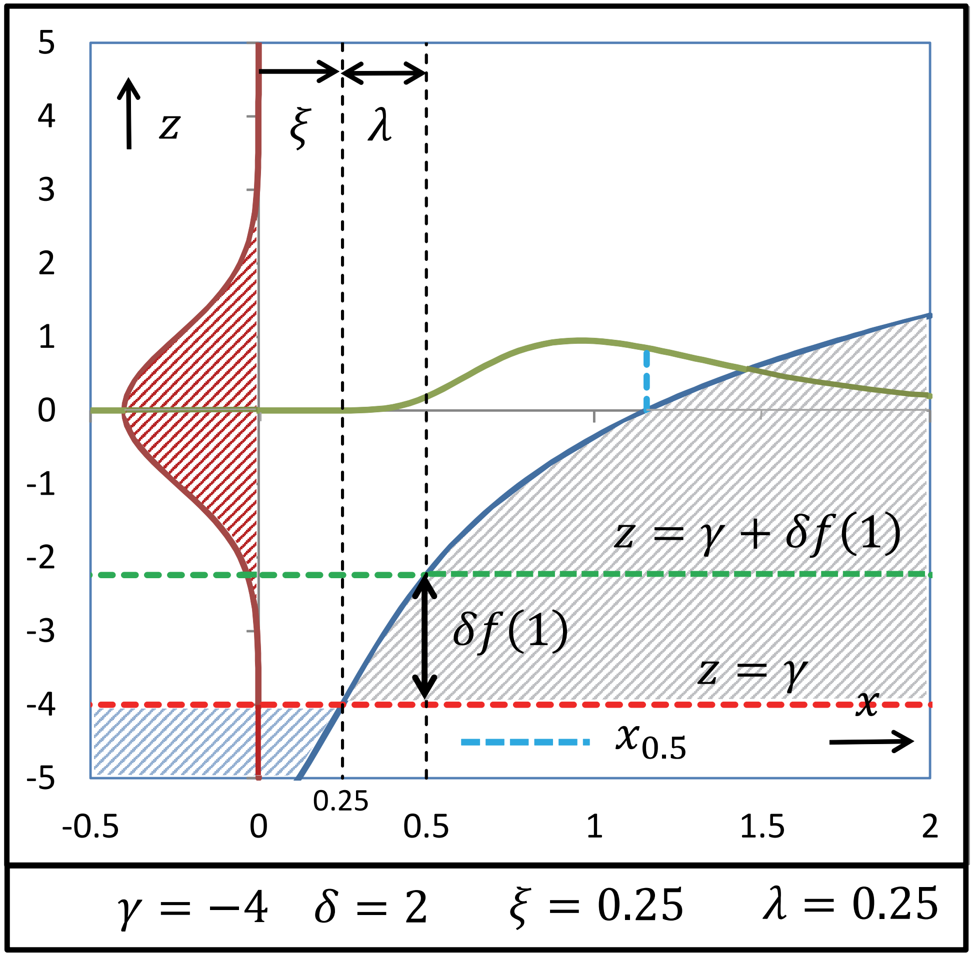
- Letting $\delta \to \infty$, but **keeping γ/δ fixed**, the median $x_{0.5}$ remains while the tangent line of $\gamma + \delta f\{(x-\xi)/\lambda\}$ converges to **vertical line at $x = x_{0.5}$.**



- S_U Pdf $p(x) \rightarrow$ **single point mass at $x_{0.5}$** as $\delta \rightarrow \infty$, **keeping γ/δ fixed.**



- Letting $\delta \downarrow 0$, but **keeping γ fixed**, the tangent line of $\gamma + \delta f\{(x-\xi)/\lambda\}$ at (ξ, γ) converges to **the horizontal line $z = \gamma$.**



- S_U Pdf $p(x) \rightarrow$ **Probability mass $\Phi(\gamma)$ at ∞** and **probability mass $1 - \Phi(\gamma)$ at ∞** as $\delta \downarrow 0$, **keeping γ fixed.**

6. SUMMARY AND CONCLUSION

- **In this geometric rediscovery** of Johnson's (1949) frequency curves we **shed light on the limiting distributions** of the three system of frequency curves **as a function of their parameters.**

- Through this discovery, **the elegance and flexibility** of **Johnson's powerful system of transformations of the normal distribution** has been demonstrated.

- **Johnson's system of transformations of the normal distribution** that **were developed in the late forties of the 20-th century**, **before the introduction of even the most primitive computers into statistical practice.**

- While **Johnson's computations were performed with the old-fashioned calculators**, **requiring copious amounts of time and effort**, these calculations **now take very little time** using modern day computational facilities.

- As such, the Johnson's (1949) frequency curves **paved the way for introduction of computation intensive methodology in statistical practice** thereby changing the game.

7. REFERENCES

Gaddum, J. H. (1945), "Lognormal distributions," *Nature*, 156(3964), 463.

Johnson, N. L. (1949), "Systems of frequency curves generated by methods of translation," *Biometrika*, 36(1/2), 149-176.

Johnson, N. L., and Kotz, S. (1985), *Distributions in Statistics*. 4 Volume Set, New York, NY: Wiley & Sons.

Kotz, S., and Johnson, N. L. (1989), *Encyclopedia of statistics*, 10 Volume Set, New York, NY: Wiley & Sons.

Kotz, S., and Johnson, N. L. (1993), *Breakthrough in statistics*. 3 Volume Set, New York, NY: Springer-Verlag.

Kumaraswamy, P. (1980), "A generalized probability density function for double-bounded random processes," *Journal of Hydrology*, 46(1-2), 79-88.

Patil, G. P. (1984), *Dictionary and classified bibliography of statistical distributions in scientific work: Continuous univariate models*, Vol. 2, Fairland, MD: International Co-operative Pub- lishing House.

Pearson, K. (1895), "Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material," *Philosophical transactions of the Royal Society of London*, 186(Part I), 343-424.

Read, C. B. (2004), "A conversation with Norman L. Johnson," *Statistical Science*, pp. 544-560.

Stuart, A., and Ord, J. K. (1994), *Kendalls advanced theory of statistics, volume 1: Distribution theory*, New York, NY: Wiley & Sons.

van Dorp, J. R., and Kotz, S. (2002), "A novel extension of the triangular distribution and its parameter estimation," *Journal of the Royal Statistical Society: Series D* (The Statistician), 51(1), 63-79.