BEYOND BETA, Other Continuous Distribution with Bounded Support and Applications

"Presentation Short Course: Beyond Beta and Applications"
***November 20th, 2018**, La Sapienza*

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

Presentation by: 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

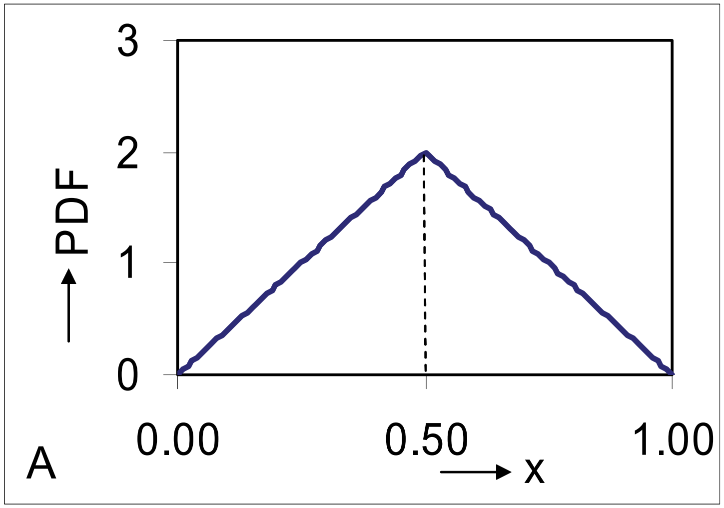
OUTLINE

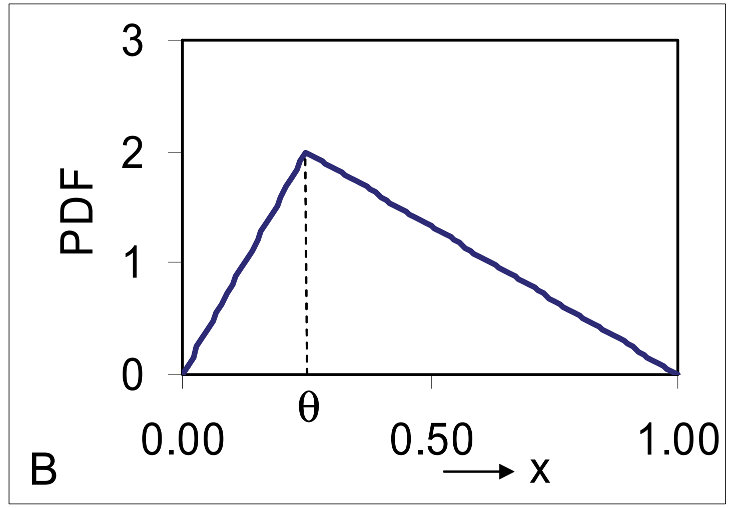
1. **Review properties of Triangular Distribution.**
2. Early extensions of the symmetric Triangular Distribution.
3. The Two-Sided Power Distribution (TSP)
4. A link between TSP and Asymmetric Laplace distributions
5. Moment Ratio Diagram Comparison
6. Mazimum Likelihood Estimation STSP distributions
7. A general framework of Two-Sided Distributions
8. Generalized Trapezoidal Distributions
9. Generalized Two-Sided Power Distributions

- One of the first records that mentions **the *continuous* uniform distribution** is the famous paper by the reverend **Thomas Bayes (1763)** (only a few years after Simpson's written records in 1757).

- One of **the earliest mentions of the triangular distributions** seems to be by **Thomas Simpson(1755, 1757)**, a colorful personality in Georgian England.

- Hence, **the continuous triangular distribution** is certainly **amongst the first *continuous* distributions** to have been noticed by investigators during the 18-th century.





- R. Schmidt (1934) was possibly the first to notice that **the symmetric triangular distribution** follows from two independent uniform random variables U_1 and U_2 on $[0, 1]$, via

$$X = \frac{U_1 + U_2}{2}.$$

- **Asymmetric standard triangular distributions** in Figure B above with support $[0, 1]$ were studied by **Ayyangar (1941)**.

- **probability density function (pdf):**

$$f(z|a,m,b)=\begin{cases}\frac{2}{b-a}\frac{z-a}{m-a}, & \text{for } a\le z\le m,\\ \frac{2}{b-a}\frac{b-z}{b-m}, & \text{for } m\le z\le b,\\ 0, & \text{elsewhere}\end{cases}$$

- **cumulative distribution function (cdf):**

$$F(z)=Pr(Z\le z|a,m,b)=\begin{cases}\frac{m-a}{b-a}\left(\frac{z-a}{z-a}\right)^2, & \text{for } a\le z\le m,\\ 1-\frac{b-m}{b-a}\left(\frac{b-z}{b-m}\right)^2 & \text{for } m\le z\le b.\end{cases}$$

- **quantile function (qf):**

$$F^{-1}(y|a,m,b,n)=\begin{cases}a+\sqrt{y(m-a)(b-a)}, & \text{for } 0\le y\le \frac{m-a}{b-a}\\ b-\sqrt{(1-y)(b-m)(b-a)}, & \text{for } \frac{m-a}{b-a}\le y\le 1\,.\end{cases}$$

- **The mean** is **the simple average** of the parameters a, m and b.

$$E[Z] = \frac{a+m+b}{3}\ .$$

- **The mode probability (MP)** to the left of the mode equals **the relative distance of the mode m to the lower bound a over the whole range $[a, b]$**.

$$Pr(Z \leq m) = \frac{m-a}{b-a} = q$$

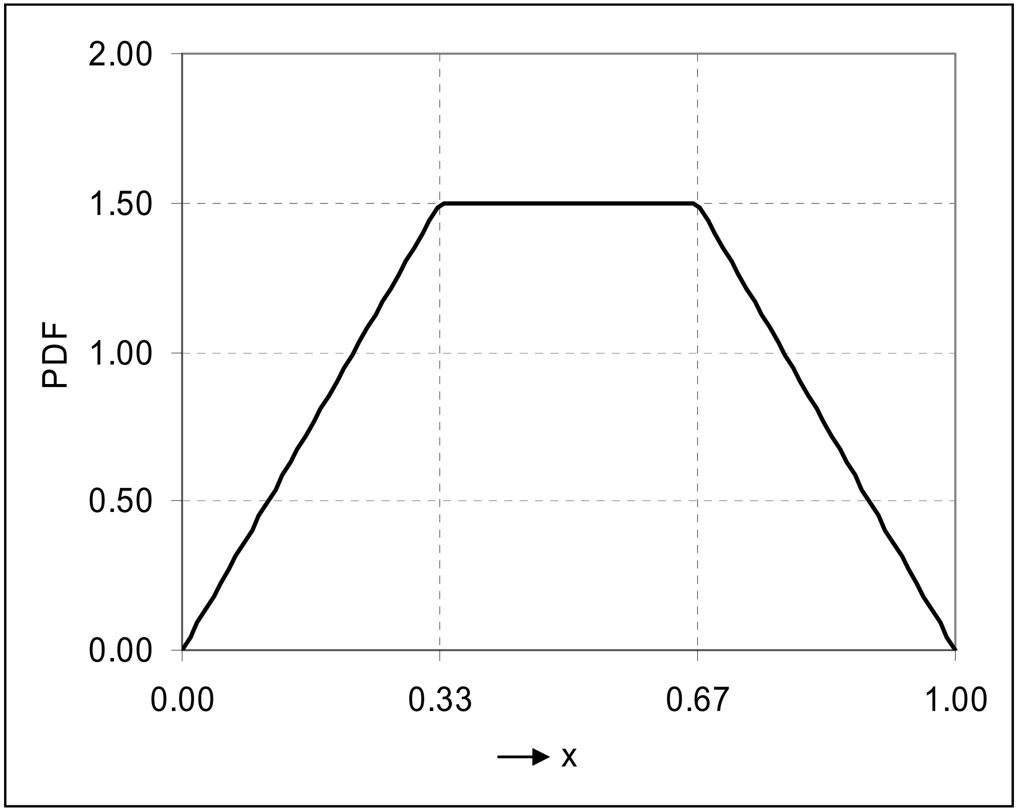
- **The variance** is **a function of that mode probability q :**

$$Var[Z] = (b-a)^2 \cdot \left\{ \frac{1 - q \times (1-q)}{18} \right\}$$

OUTLINE

1. Review properties of Triangular Distribution.
2. **Early extensions of the Triangular Distribution.**
3. The Two-Sided Power Distribution (TSP)
4. A link between TSP and Asymmetric Laplace distributions
5. Moment Ratio Diagram Comparison
6. Mazimum Likelihood Estimation STSP distributions
7. A general framework of Two-Sided Distributions
8. Generalized Trapezoidal Distributions
9. Generalized Two-Sided Power Distributions

THE SYMMETRIC TRAPEZOIDAL DISTRIBUTION



$$X = \frac{1}{3}U_1 + \frac{2}{3}U_2$$

A finite mixture of Uniform Distributions may thus be interpreted as **an extension of the symmetric Triangular distribution with support [0, 1].**

$$Y=\sum_{i=1}^{m} w_i U_i,\ \sum_{i=1}^{m} w_i = 1,\ w_i \geq 0,$$

$$Pr(Y \leq y)=\sum_{v_1=0}^{1}\cdots\sum_{v_m=0}^{1}(-1)^{\sum_{i=1}^{m} v_i}\left\{\frac{(y-\sum_{i=1}^{m} w_i v_i)^m}{m!\prod_{i=1}^{m} w_i}\right\}\mathbf{1}_{[0,\infty)}(y-\sum_{i=1}^{n} w_i v_i).$$

- First derived by **Mitra (1971)**. Later by **Barrow and Smith (1979)**. Their proofs are geared towards mathematically oriented readers, are very concise and somewhat difficult to follow.

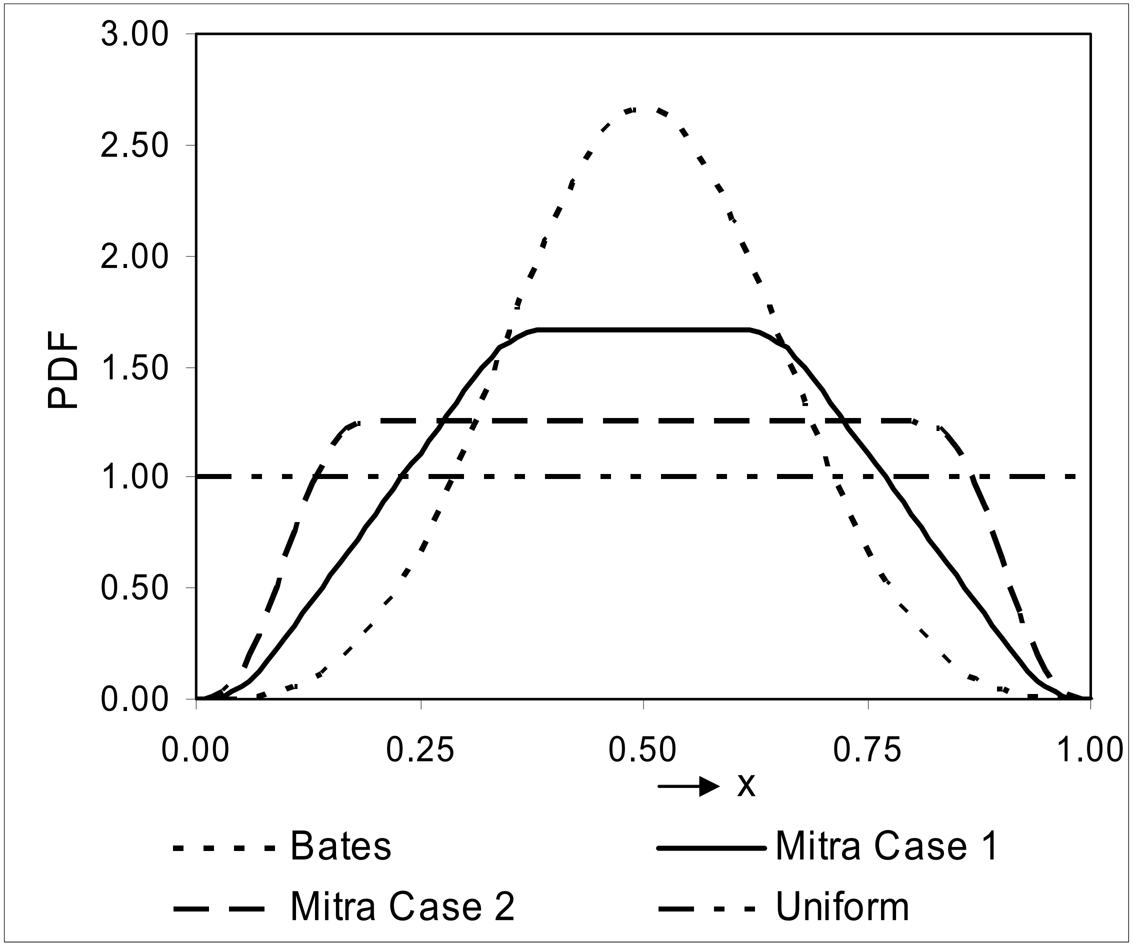


Figure 2.8. Examples of the pdf's of (2.0) for $n = 4$. A: $w_i = \frac{1}{4}, i = 1, \ldots, 4$ (Bates); B: $w_1 = 0.6, w_2 = 0.3, w_3 = 0.075,\ w_4 = 0.025$ (Mitra Case 1); C: $w_1 = 0.8, w_2 = 0.1, w_3 = 0.075,\ w_4 = 0.025$ (Mitra Case 2); D: $w_1 = 1,\ w_i = 0,\ i = 2, 3, 4$ (Uniform).

- Based on the time honored **inclusion-exclusion principle**

$$Pr\Big\{\bigcup_{i=1}^{m} A_i\Big\} = \sum_{i=1}^{m} Pr(A_i) - \sum\sum_{i<j} Pr\{A_i \cap A_j\} +$$
$$\sum\sum\sum_{i<j<k} Pr\{A_i \cap A_j \cap A_k\} - \ldots + (-1)^m Pr\Big\{\bigcap_{i=1}^{m} A_i\Big\},$$

for arbitrary events $A_1, \ldots, A_n$ (not necessarily disjoint) (see, e.g., Feller (1990)).

- The geometric nature of the proof allows for **an efficient algorithm for evaluation of the cdf of Y** useful for its **application in Monte Carlo based uncertainty analyses.**

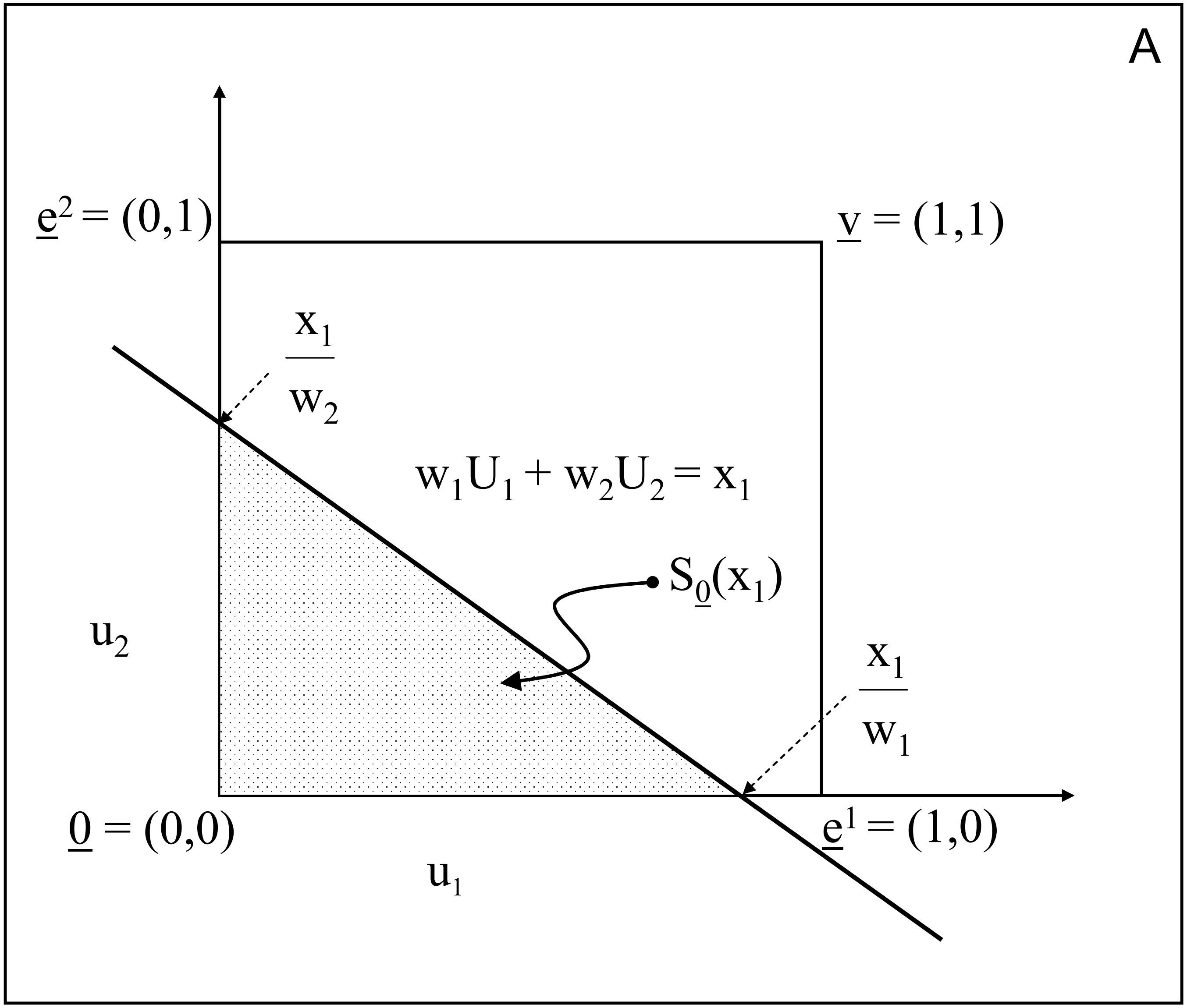
- Let $C^m = \{\underline{u} \,|\, 0 \le u_i \le 1\}$ **be the unit hyper cube in** $\mathbb{R}^m$. Let $\underline{v} = (v_1, \ldots, v_m)$, $v_i \in \{0, 1\}$ be a vertex **(corner-point)** of the unit hyper cube C^m and define **the simplex $S_{\underline{v}}(y)$ at vertex $\underline{v}$** as

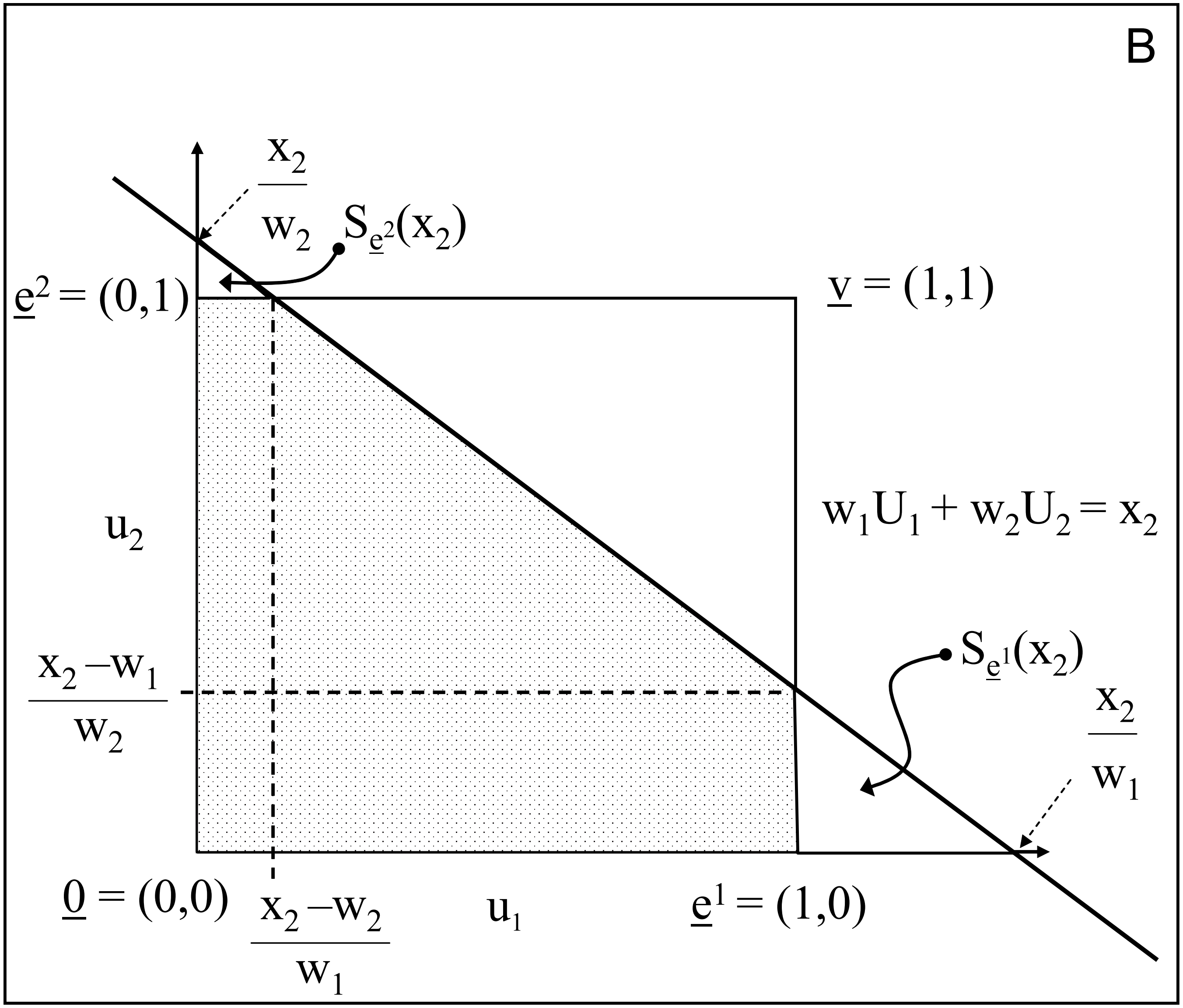
$$S_{\underline{v}}(y) = \{\, \underline{u} \mid \sum_{i=1}^{m} w_i u_i \le y, u_i \ge v_i, i = 1, \ldots, m \,\},$$

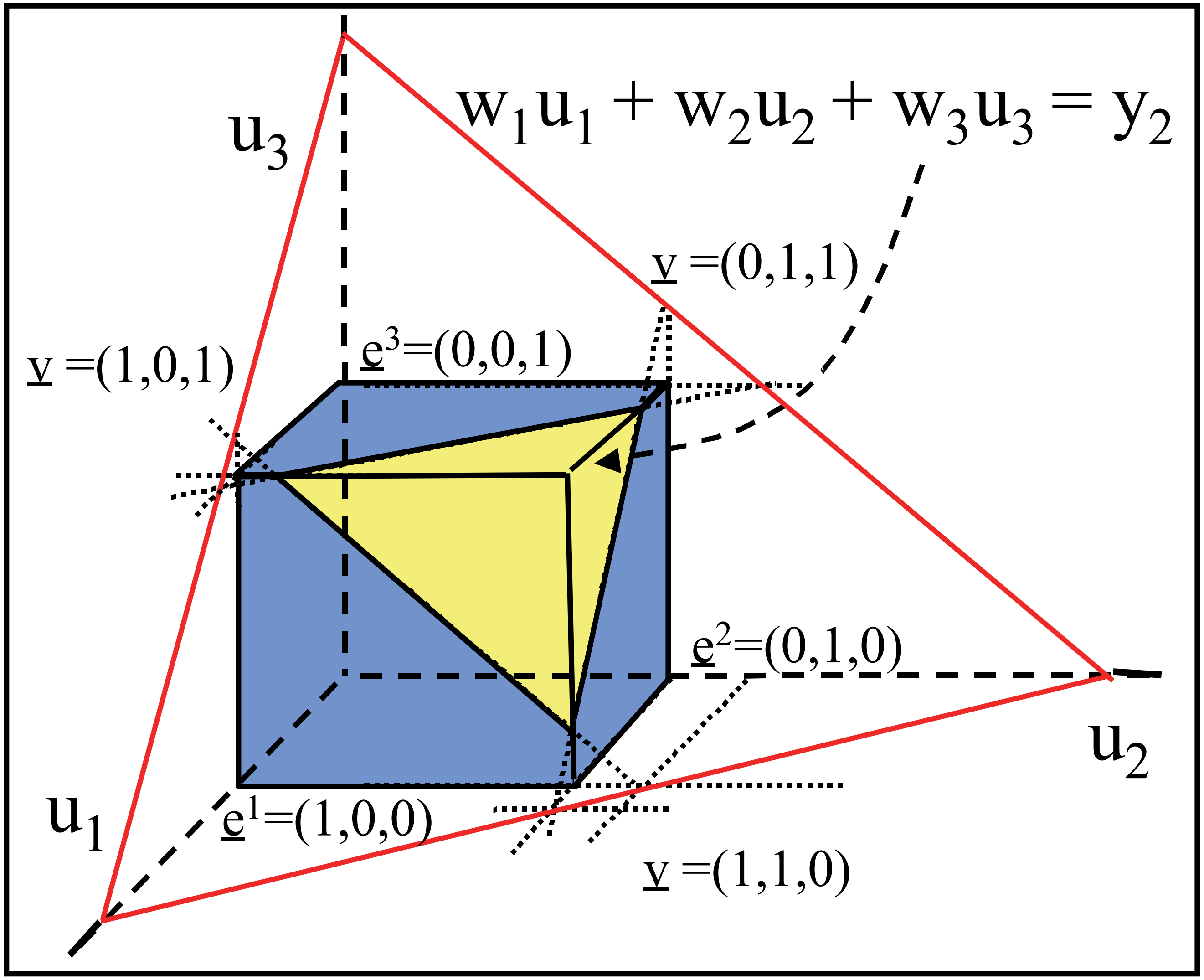
where $w_i \ge 0$, $\sum_{i=1}^{m} w_i = 1$.

- **The hyper-volume of the simplex $S_{\underline{v}}(y)$** denoted by $V\{S_{\underline{v}}(y)\}$ is given by:

$$V\{S_{\underline{v}}(y)\} = \frac{(y - \sum_{i=1}^{m} w_i v_i)^m}{m! \prod_{i=1}^{m} w_i} \cdot 1_{[0,\infty)}(y - \sum_{i=1}^{n} w_i v_i).$$


A
e² = (0,1)
v = (1,1)
x1 / w2
w1U1 + w2U2 = x1
S0(x1)
u2
x1 / w1
0 = (0,0)
e¹ = (1,0)
u1


B
x2
w2
Se2(x2)
e2 = (0,1)
v = (1,1)
u2
w1U1 + w2U2 = x2
Se1(x2)
x2 −w1
w2
x2
w1
0 = (0,0)
x2 −w2
w1
u1
e1 = (1,0)



Evaluating $G(y_2)$ for $m = 3$.

- Trapezoidal distributions have been advocated in risk analysis problems **by Pouliquen (1970)**. **Its pdf is given by :**

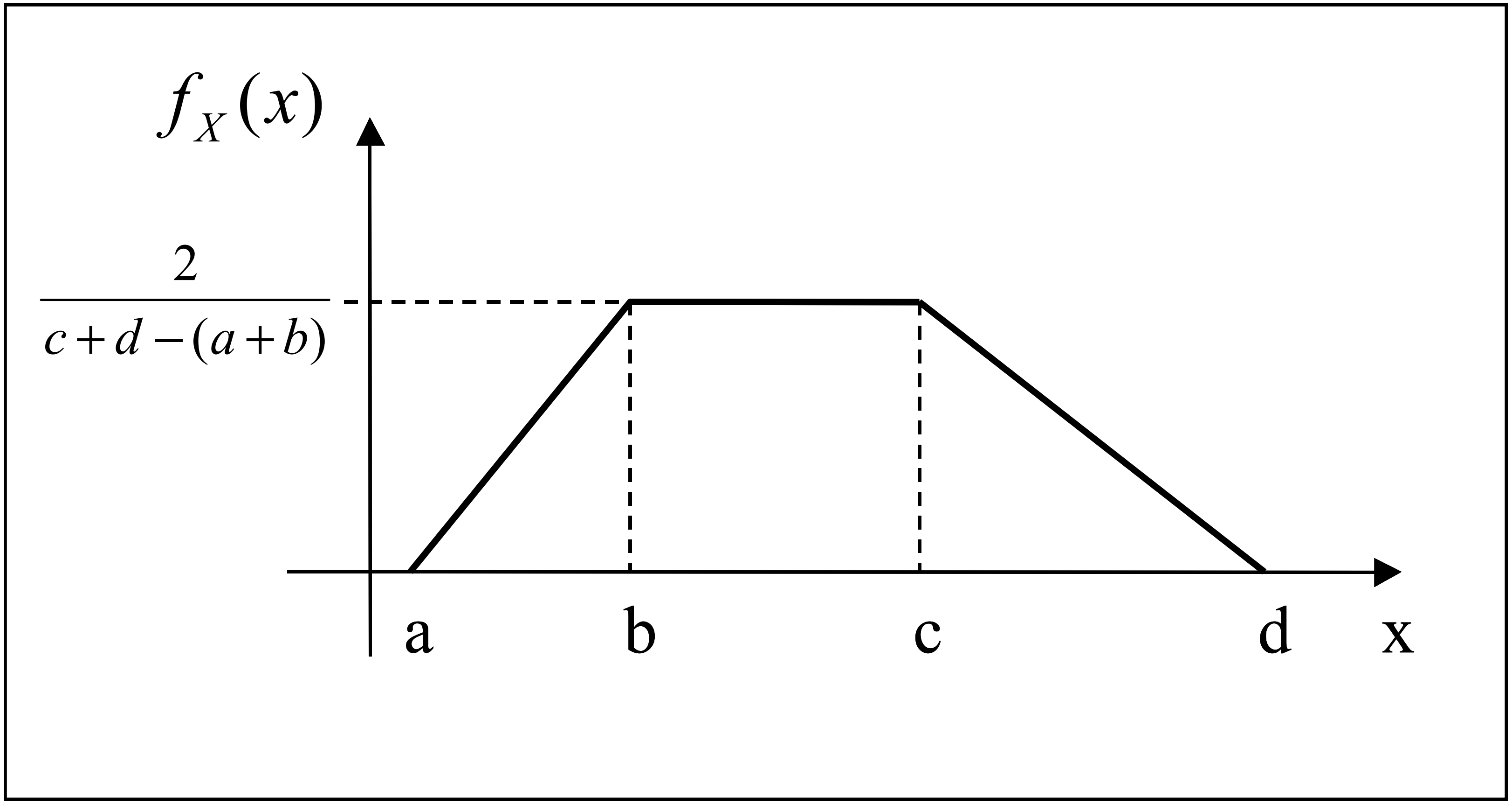
$$f_X(x|a,b,c,d) = \mathcal{C}(a,b,c,d) \times \begin{cases} \frac{x-a}{b-a}, & \text{for } a \le x < b \\ 1, & \text{for } b \le x < c \\ \frac{d-x}{d-c}, & \text{for } c \le x < d \end{cases}$$

where $a < b < c < d$ and $\mathcal{C}(a,b,c,d) = 2\{c+d-(a+b)\}^{-1}$.

- **Cumulative distribution function (cdf):**

$$F_X(x|a,b,c,d) = \begin{cases} \mathcal{C}(a,b,c,d)\frac{(b-a)}{2}\left(\frac{x-a}{b-a}\right)^2, & \text{for } a \le x < b \\ \mathcal{C}(a,b,c,d)(x-(b+a)/2), & \text{for } b \le x < c \\ 1-\mathcal{C}(a,b,c,d)\frac{d-c}{2}\left(\frac{d-x}{d-c}\right)^2, & \text{for } c \le x < d. \end{cases}$$

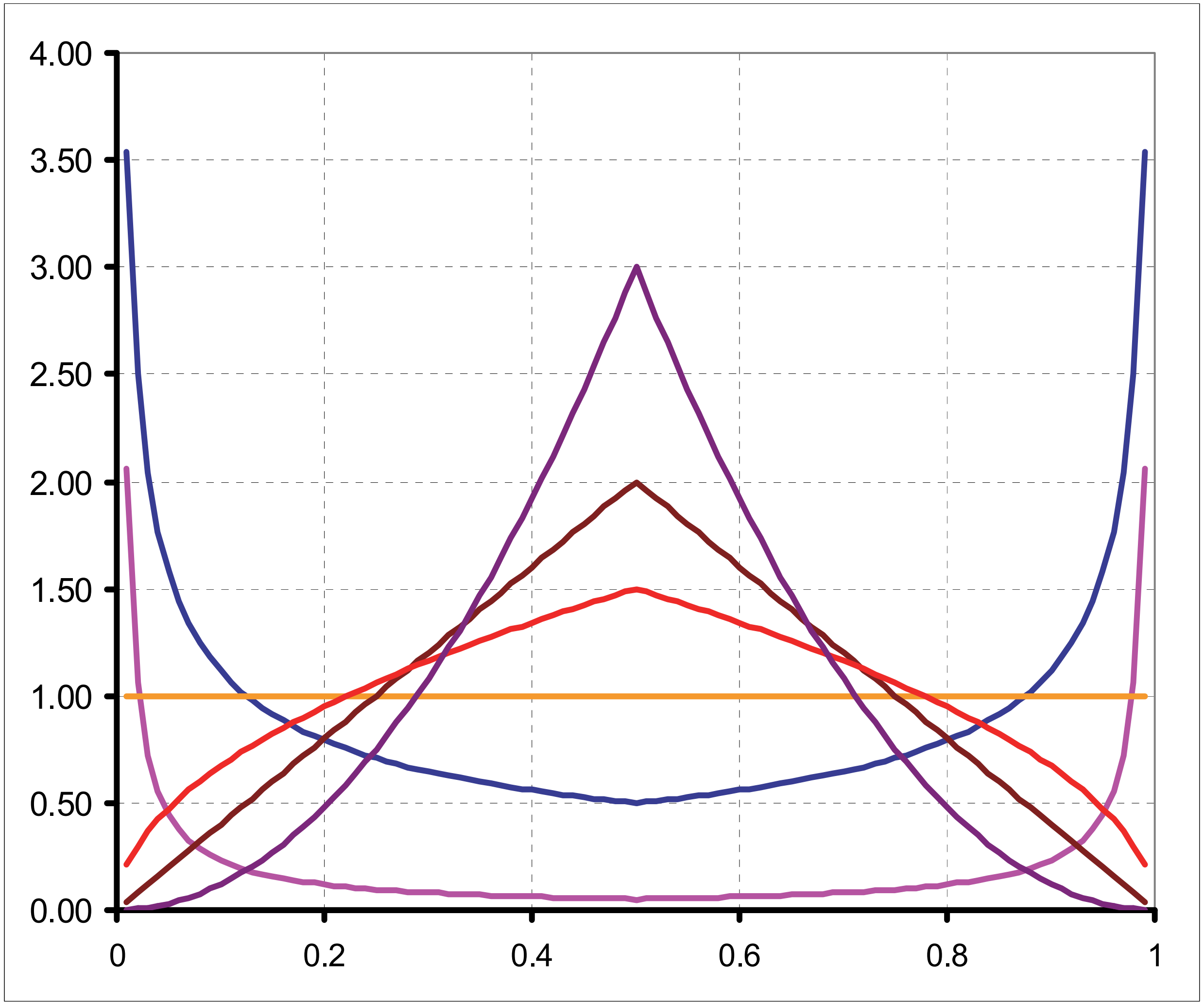
,



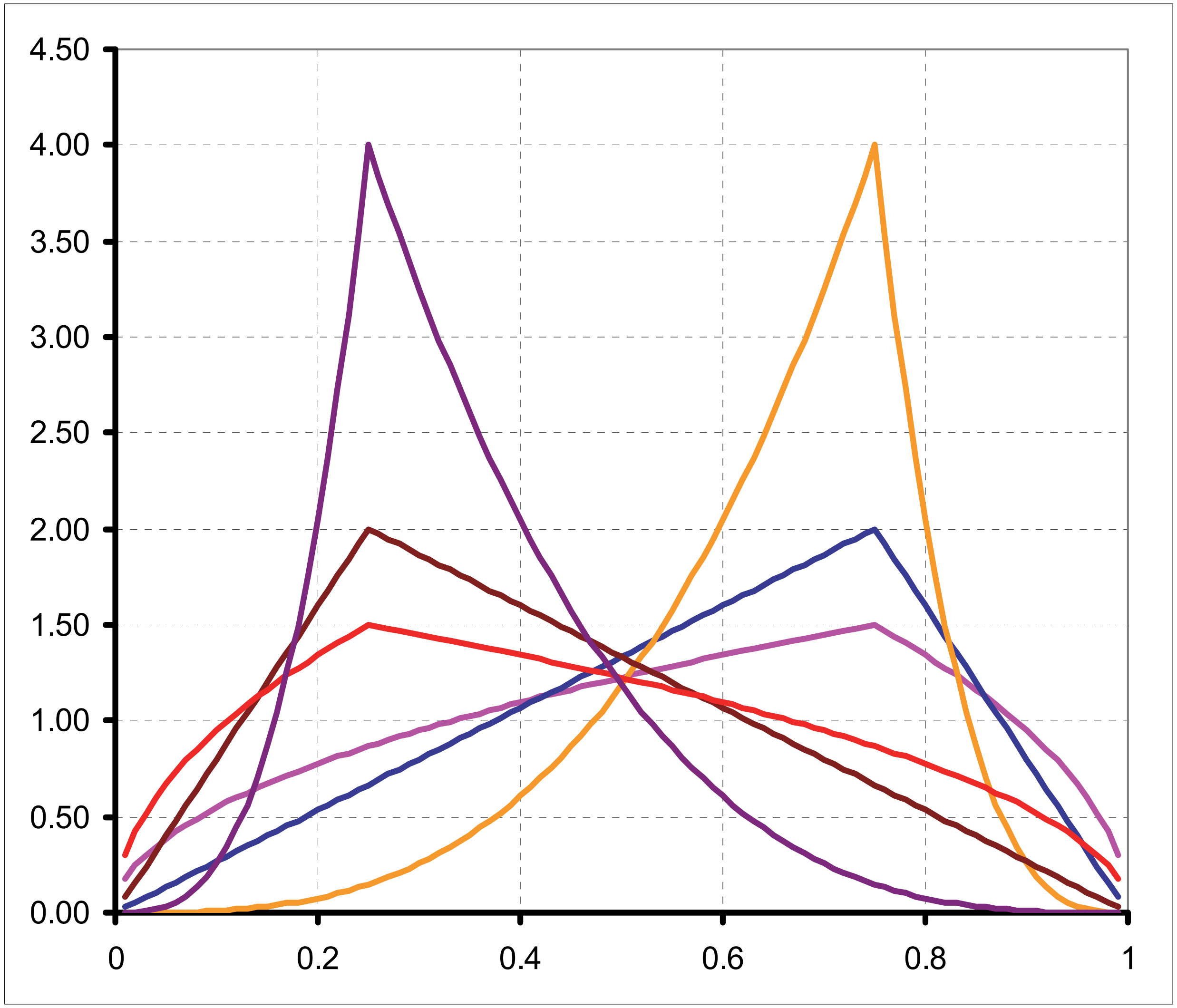
Probability density function of a **Trapezoidal Distribution**.

OUTLINE

1. Review properties of Triangular Distribution.
2. Early extensions of the Triangular Distribution.
3. **The Two-Sided Power Distribution (TSP)**
4. A link between TSP and Asymmetric Laplace distributions
5. Moment Ratio Diagram Comparison
6. Mazimum Likelihood Estimation STSP distributions
7. A general framework of Two-Sided Distributions
8. Generalized Trapezoidal Distributions
9. Generalized Two-Sided Power Distributions



Symmetric Two-Sided Power Distribution Examples



Asymmetric Two-Sided Power Distribution Examples

- **probability density function (pdf)** for $a \leq m \leq b,\ n > 0$:

$$f_Z(x|a,m,b,n) = \begin{cases} \frac{n}{b-a}\left(\frac{z-a}{m-a}\right)^{n-1}, & \text{for } a < z \leq m \\ \frac{n}{b-a}\left(\frac{b-z}{b-m}\right)^{n-1}, & \text{for } m \leq z < b \end{cases}$$

- **cumulative distribution function (pdf):**

$$F_Z(x|a,m,b,n) = \begin{cases} \frac{m-a}{b-a}\left(\frac{z-a}{m-a}\right)^{n}, & \text{for } a \leq z \leq m \\ 1 - \frac{b-m}{b-a}\left(\frac{b-z}{b-m}\right)^{n}, & \text{for } m \leq z \leq b\,. \end{cases}$$

- **quantile function (pdf),** $q = (m-a)/b-a)$:

$$F_Z^{-1}(y|a,m,b,n) = \begin{cases} a + \sqrt[n]{y(m-a)^{n-1}(b-a)}, & \text{for } 0 \leq y \leq q \\ b - \sqrt[n]{(1-y)(b-m)^{n-1}(b-a)}, & \text{for } q \leq y \leq 1\,. \end{cases}$$

For $\boldsymbol{n = 2}$ the above expressions reduce to the triangular equivalents.

- The mean is now **a weighted average** of the parameters a, m and b.

$$E[Z] = \frac{a + (n-1)m + b}{n+1}$$

- **The mode probability (MP)** to the left of the mode equals **the relative distance of the mode m to the lower bound a over the range $[a, b]$, regardless of the value of the power parameter n.**

$$Pr(Z \leq m) = \frac{m-a}{b-a} = q$$

- **The variance** is **a function of the mode probability q above :**

$$Var[Z] = (b-a)^2 \cdot \left\{ \frac{n - 2(n-1) \times q(1-q)}{(n+2)(n+1)^2} \right\}$$

- By substituting $n = 1$ above we obtain the equivalent expressions for **the uniform distribution.**

OUTLINE

1. Review properties of Triangular Distribution.
2. Early extensions of the Triangular Distribution.
3. The Two-Sided Power Distribution (TSP).
4. **A link between TSP and Asymmetric Laplace distributions**
5. Moment Ratio Diagram Comparison
6. Mazimum Likelihood Estimation STSP distributions
7. A general framework of Two-Sided Distributions
8. Generalized Trapezoidal Distributions
9. Generalized Two-Sided Power Distributions

- Suppose **a lower percentile a_p** and **upper percentile b_r** and **the mode m** are specified and we wish to solve for **the lower bound a** and **upper bound b,** where $a < a_p < m < b_r < b$.

- Recalling **the mode probability** for TSP distributions:

$$Pr(Z < m|a, m, b, n) \equiv q = (m - a)/(b - a)$$

one obtains with $p = 1 - q$ after some algebraic manipulations :

$$a \equiv a(q|n) = \frac{a_p - m\sqrt[n]{p/q}}{1 - \sqrt[n]{p/q}} < \frac{a_p - a_p\sqrt[n]{p/q}}{1 - \sqrt[n]{p/q}} = a_p.$$

$$b \equiv b(q|n) = \frac{b_r - m\sqrt[n]{\frac{1-r}{1-q}}}{1 - \sqrt[n]{\frac{1-r}{1-q}}} > \frac{b_r - b_r\sqrt[n]{\frac{1-r}{1-q}}}{1 - \sqrt[n]{\frac{1-r}{1-q}}} = b_r.$$

- Substituting $a(q|n)$ and $b(q|n)$ back in **the mode probability equation:**

$$q = (m - a)/(b - a)$$

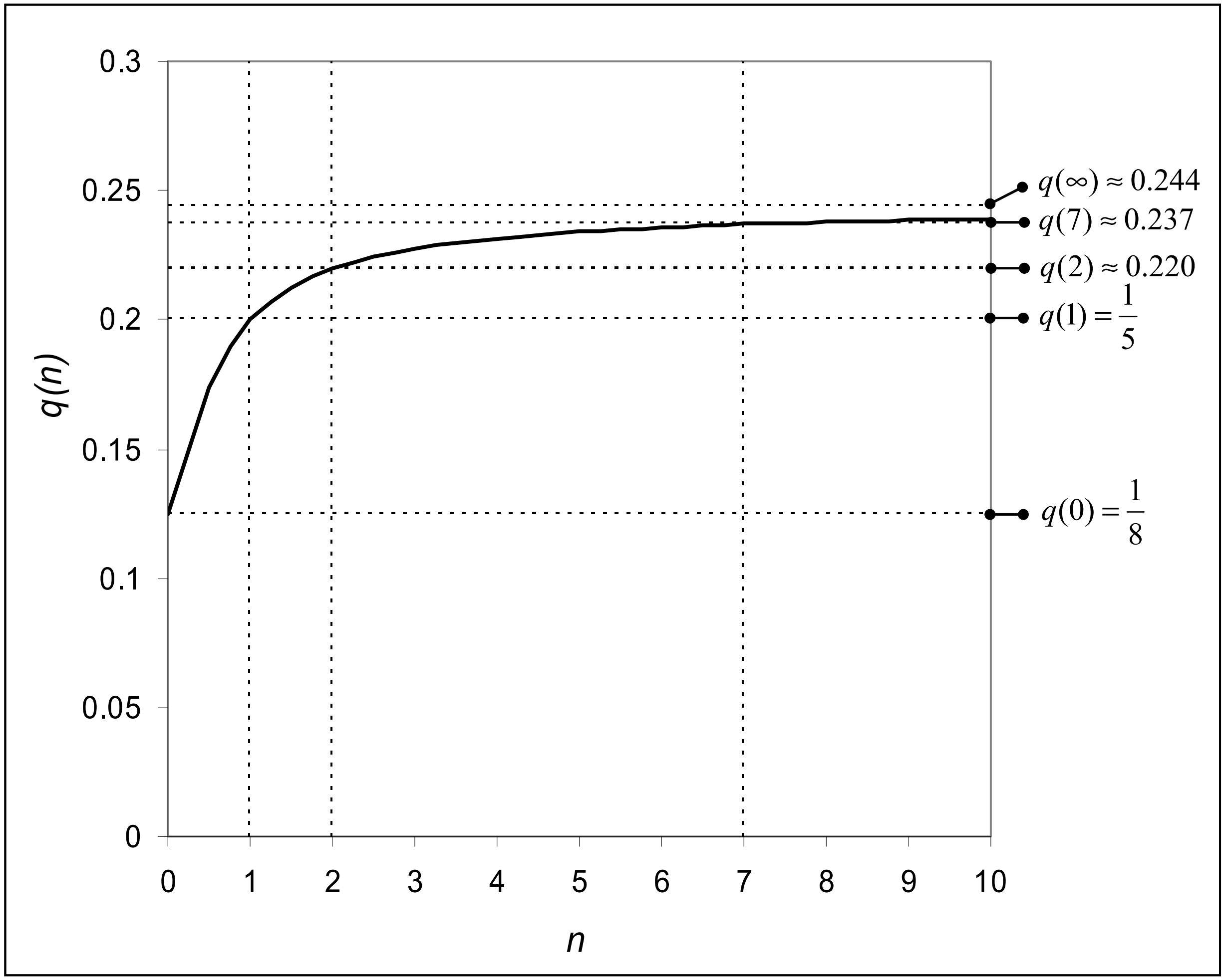
one obtains the equation

$$q = \frac{m - a(q|n)}{b(q|n) - a(q|n)} \equiv g(q|n)$$

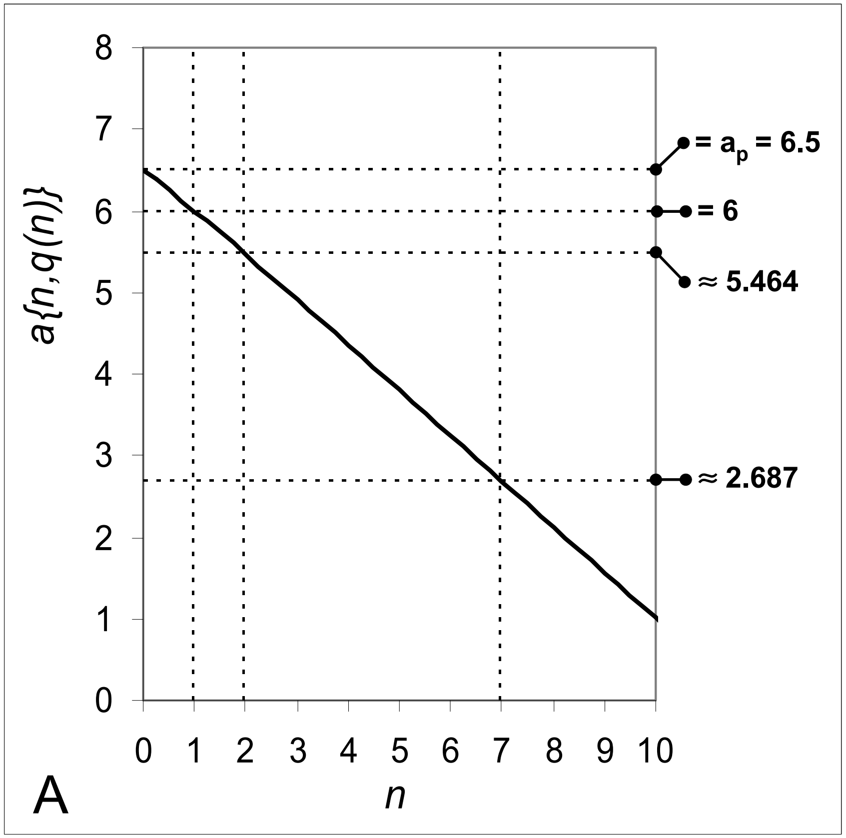
where

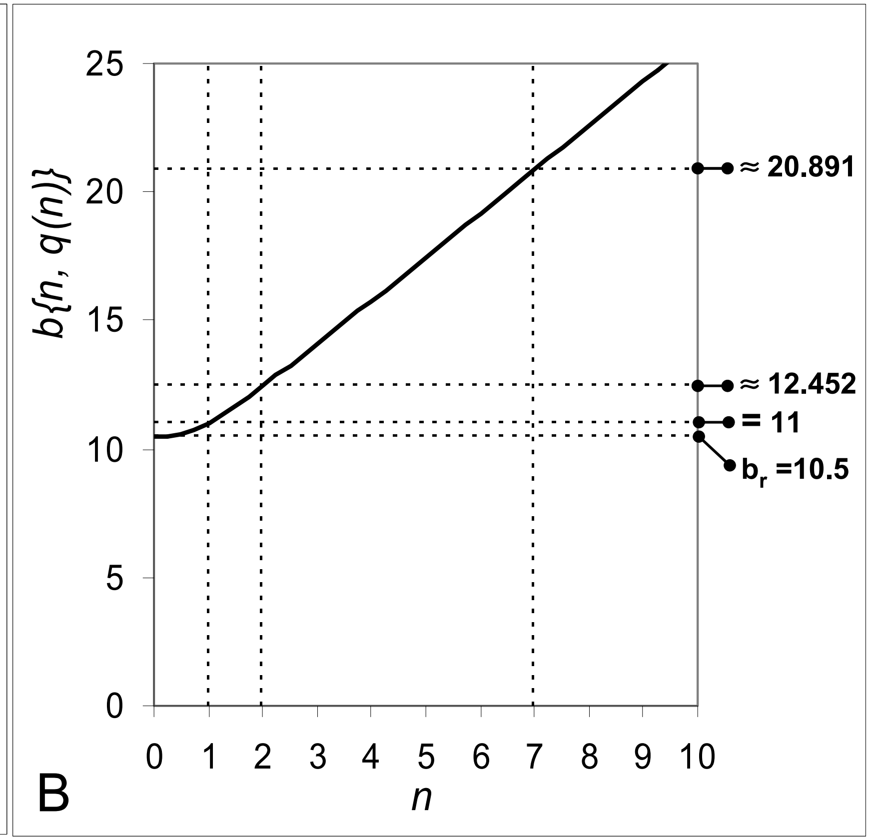
$$g(q|n) \equiv \frac{(m - a_p)\left(1 - \sqrt[n]{\frac{1-r}{1-q}}\right)}{(b_r - m)\left(1 - \sqrt[n]{\frac{p}{q}}\right) + (m - a_p)\left(1 - \sqrt[n]{\frac{1-r}{1-q}}\right)}$$

- The equation above defines **a continuous implicit function $q(n)$ with domain $n > 0$,** where **$q(n)$ is the probability mass to the left of the mode m** for a **given value of the shape/power parameter n**.



Graph of **the implicit function $q(n)$** for the case $p = 0.10$, $r = 1 - p = 0.90$, $a_p = 6.5$, $m = 7$ and $b_r = 10.5$.





The **lower bound function $a\{q(n)|n\}$** and **the upper bound function $b\{q(n)|n\}$** where $q(n)$ is the implicit function depicted in the previous figure for the case $p = 0.10,\ r = 0.90,\ a_p = 6.5, m = 7$ and $b_r = 10.5$.

- Note that, $q(n) \rightarrow q(\infty)$ as $n \rightarrow \infty$, where $q(\infty)$ is a fixed value.
- Note that, $a\{q(n)|n\} \rightarrow -\infty$ as $n \rightarrow \infty$.
- Note that, $b\{q(n)|n\} \rightarrow \infty$ as $n \rightarrow \infty$.
- Introducing **the relative distance δ of the mode m to the lower quantile a_p** over **the range from a_p to the upper quantile b_r**,

$$\delta = \frac{m - a_p}{b_r - a_p}$$

(and thus $\delta \in (0, 1)$,} it can be shown that **$q(\infty)$ is the unique solution** of the equation

$$\frac{q(\infty)}{\delta} Log\left\{\frac{q(\infty)}{p}\right\} = \frac{1 - q(\infty)}{1 - \delta} Log\left\{\frac{1 - q(\infty)}{1 - r}\right\}.$$

- In addition, the TSP distribution with parameters **$a\{q(n)|n\}, m,$ $b\{q(n)|n\}$** and **n** converges as $n \to \infty$ (in distribution) to the pdf

$$f_X(x|a_p,m,b_r) = \begin{cases} q(\infty)\mathcal{A}Exp\Big\{-\mathcal{A}(m-x)\Big\} & x \leq m \\ \{1-q(\infty)\}\mathcal{B}Exp\Big\{-\mathcal{B}(x-m)\Big\} & x > m, \end{cases}$$

where the coefficients $\mathcal{A}$ and $\mathcal{B}$ are

$$\mathcal{A} = \frac{Log\Big\{\frac{q(\infty)}{p}\Big\}}{m-a_p} \text{ and } \mathcal{B} = \frac{Log\Big\{\frac{1-q(\infty)}{1-r}\Big\}}{b_r-m}.$$

- Pdf $f_X(x|a_p,m,b_r)$ is a **reparameterized asymmetric Laplace distribution** with **lower and upper quantile a_p and b_r and mode m**.

Kotz, S. and Van Dorp, J.R. (2005). A Link between Two-Sided Power and Asymmetric Laplace Distributions: with Applications to Mean and Variance Approximations, *Statistics and Probability Letters*

OUTLINE

1. Review properties of Triangular Distribution.
2. Early extensions of the Triangular Distribution.
3. The Two-Sided Power Distribution (TSP).
4. A link between TSP and Asymmetric Laplace distributions
5. **Moment Ratio Diagram Comparison**
6. Mazimum Likelihood Estimation STSP distributions
7. A general framework of Two-Sided Distributions
8. Generalized Trapezoidal Distributions
9. Generalized Two-Sided Power Distributions

- Moment ratio plots provide a **useful visual assessment** of **the skewness** and **the kurtosis. Their coverage** is **indicative of the flexibility** of **a particular family of asymmetric distributions.**

- The classical form of the diagram shows the values of the ratio

$$\beta_1 = \frac{E^2[(X-E[X])^3]}{E^3[(X-E[X])^2]} = \frac{\mu_3^2}{\mu_2^3}, \ \beta_2 = \frac{E[(X-E[X])^4]}{E^2[(X-E[X])^2]} = \frac{\mu_4}{\mu_2^2},$$

where $\boldsymbol{\beta_1}$ **is plotted on the abscissa** and $\boldsymbol{\beta_2}$ **on the ordinate**.

- This diagram possesses **a disadvantage** that **the sign of** $\boldsymbol{\mu_3}$ (indicating left skewness or right skewness) **disappears.**

- A moment ratio diagram that retains this information is a plot with $\sqrt{\beta_1}$ **on the abscissa** and $\boldsymbol{\beta_2}$ **on the ordinate**, with the convention that $\sqrt{\beta_1}$ **retains the sign of** $\boldsymbol{\mu_3}$ (See, e.g., Kotz and Johnson (1985)).

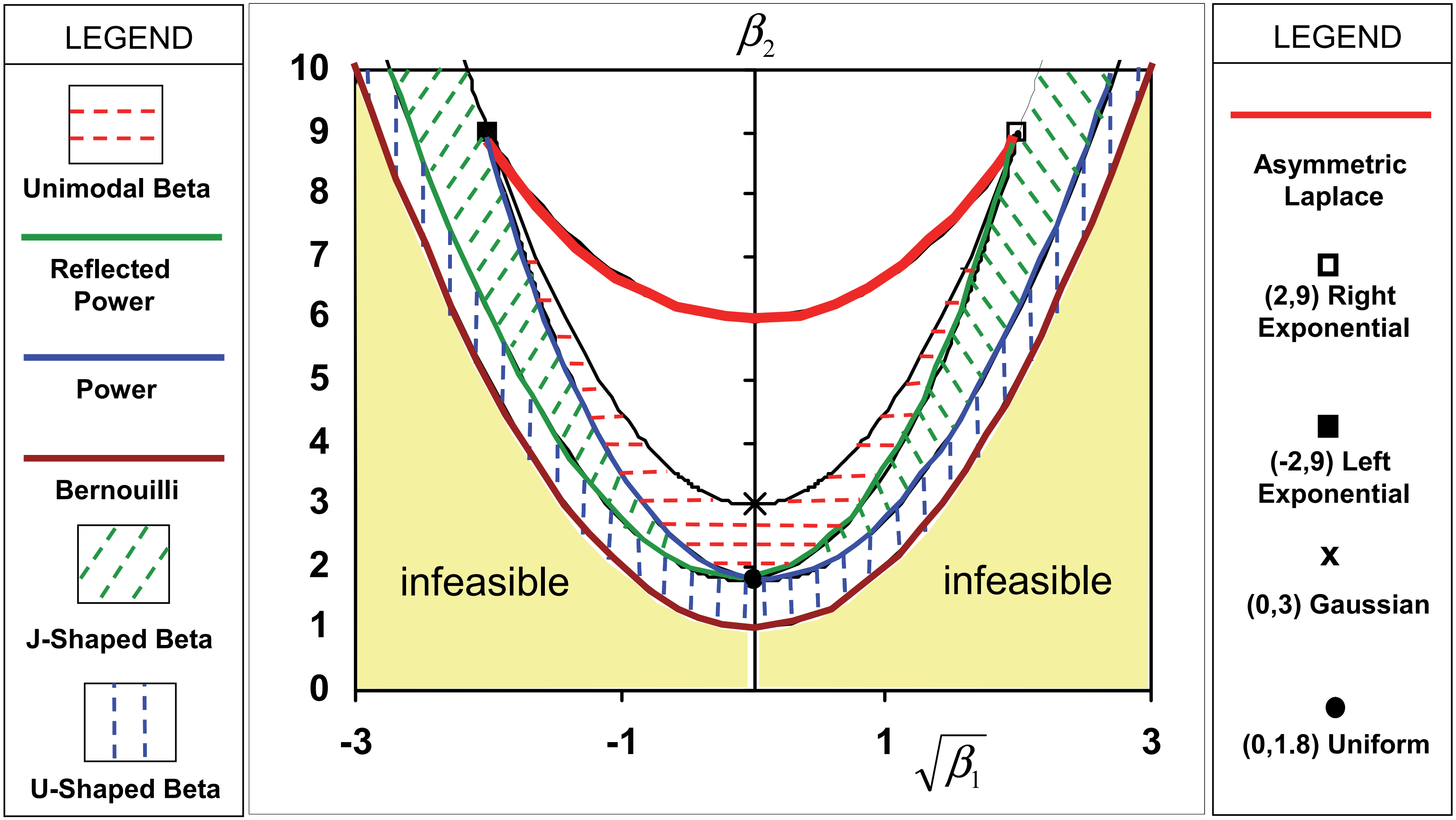
- Values for $\sqrt{\beta_1}$ and β_2 for **Standard TSP distributions** (with support $[0,1]$, mode θ and shape/power parameter n) can be calculated using the general expression for their moments **around the origin** $\boldsymbol{\mu'_k = E[X^k]}$, $k = 1, \ldots, 4$, given by

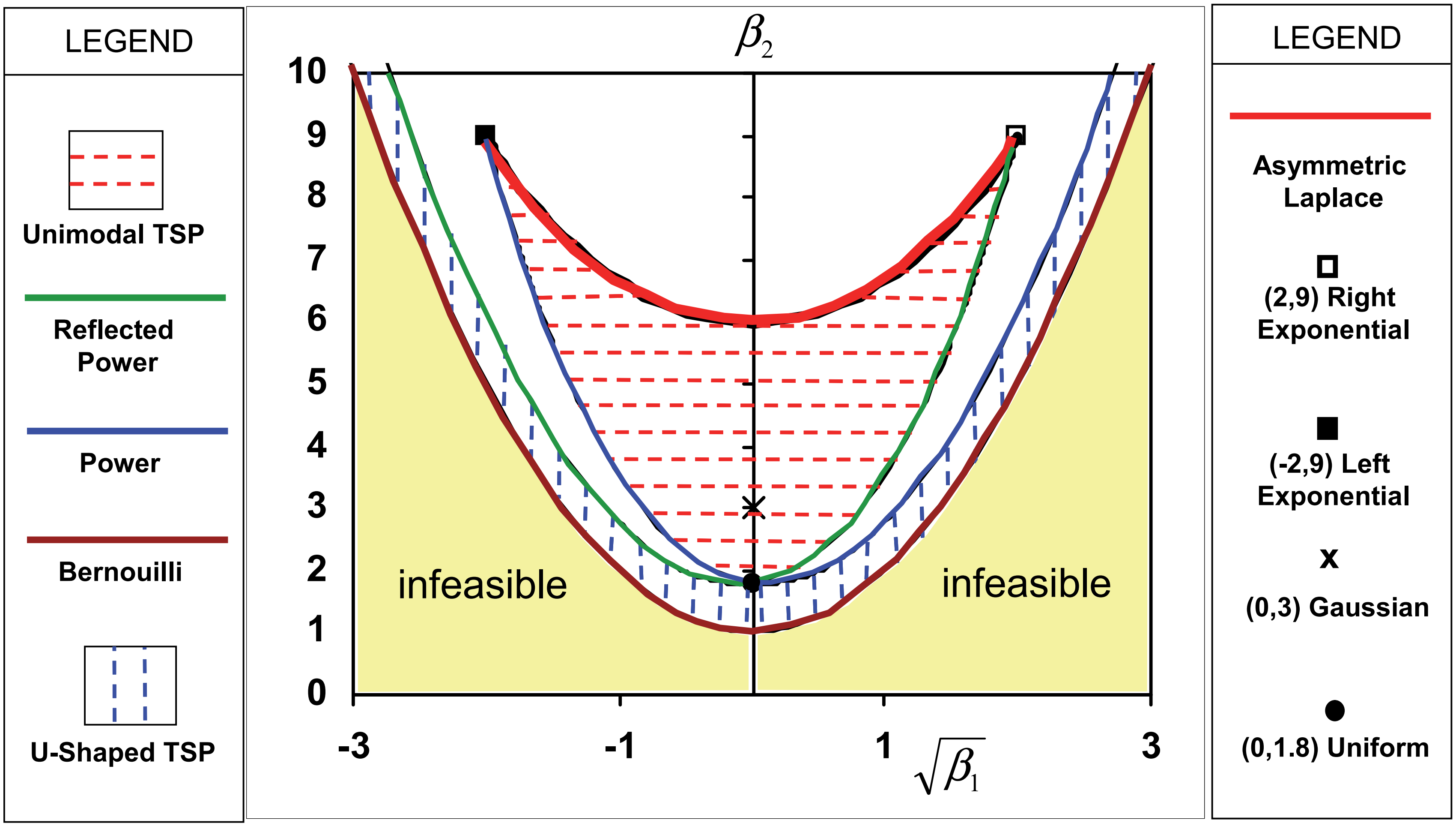
$$\mu'_k = E[X^k] = \frac{n\theta^{k+1}}{n+k} - \sum_{i=0}^{k} \binom{k}{k-i} \frac{n(\theta-1)^{i+1}}{n+i}.$$

and their relationship with **central moments** $\boldsymbol{\mu_k = E[(X-\mu_1)^k]}$, $\boldsymbol{k = 2, 3, 4}$ given by

$$\begin{cases} \mu_2 = \mu'_2 - {\mu'_1}^2 \\ \mu_3 = \mu'_3 - 3\mu'_2\mu'_1 + 2{\mu'_1}^3 \\ \mu_4 = \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2{\mu'_1}^2 - 3{\mu'_1}^4 \end{cases}$$

(see, e.g., Stuart and Ord (1994)).


LEGEND
Unimodal Beta
Reflected Power
Power
Bernouilli
J-Shaped Beta
U-Shaped Beta
β_2
10
9
8
7
6
5
4
3
2
1
0
-3
-1
1
3
$\sqrt{\beta_1}$
infeasible
infeasible
LEGEND
Asymmetric Laplace
(2,9) Right Exponential
(-2,9) Left Exponential
x
(0,3) Gaussian
(0,1.8) Uniform



LEGEND
Unimodal TSP
Reflected Power
Power
Bernouilli
U-Shaped TSP
β_2
10
9
8
7
6
5
4
3
2
1
0
-3
-1
1
3
$\sqrt{\beta_1}$
infeasible
infeasible
LEGEND
Asymmetric Laplace
(2,9) Right Exponential
(-2,9) Left Exponential
x
(0,3) Gaussian
(0,1.8) Uniform

OUTLINE

1. Review properties of Triangular Distribution.
2. Early extensions of the Triangular Distribution.
3. The Two-Sided Power Distribution (TSP).
4. A link between TSP and Asymmetric Laplace distributions
5. Moment Ratio Diagram Comparison
6. **Mazimum Likelihood Estimation STSP distributions**
7. A general framework of Two-Sided Distributions
8. Generalized Trapezoidal Distributions
9. Generalized Two-Sided Power Distributions

- **Standard TSP (STSP) distribution** pdf with support $[0, 1]$:

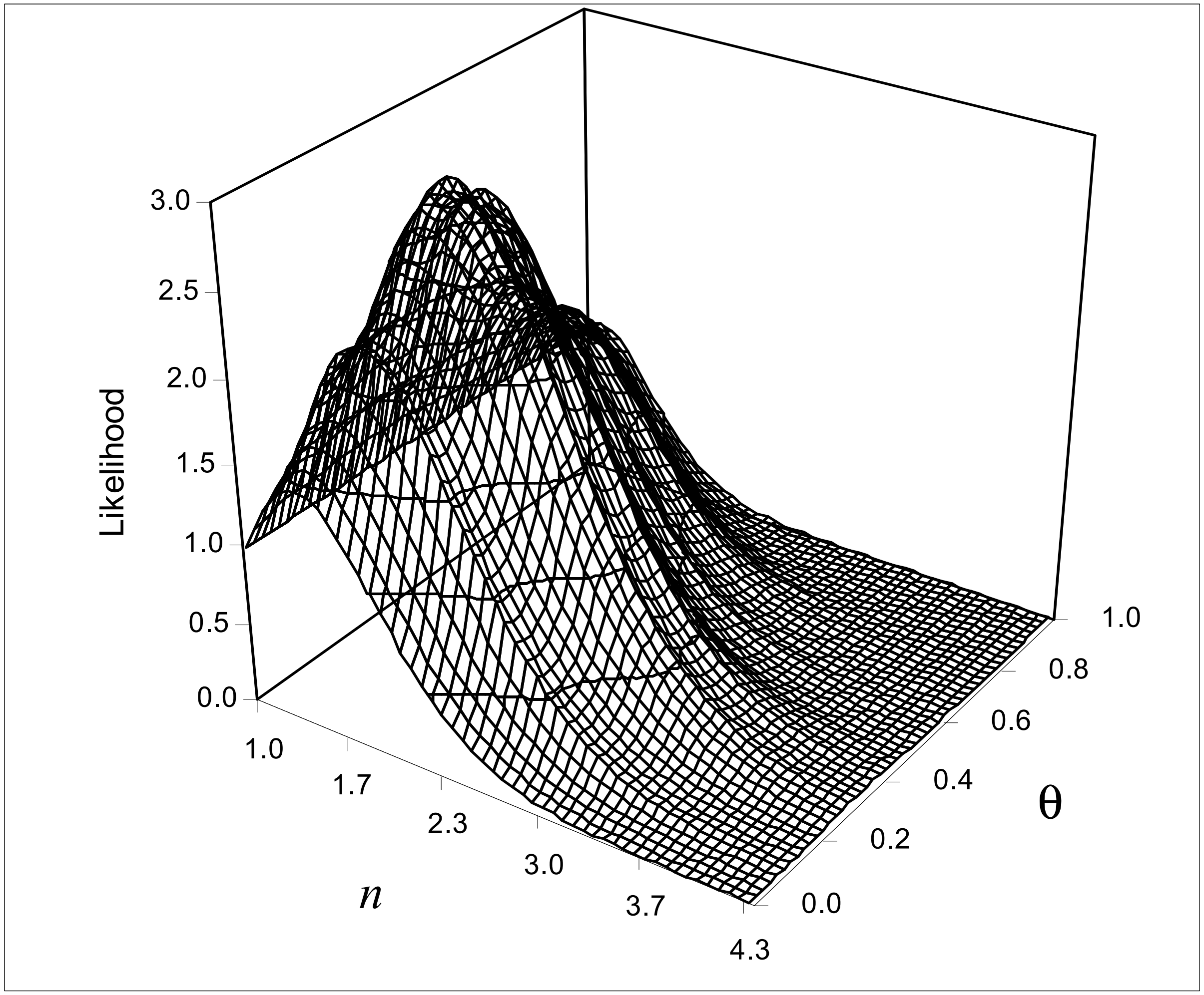
$$f(x|\theta,n) = \begin{cases} n\left(\frac{x}{\theta}\right)^{n-1}, & \text{for } 0 \leq x \leq \theta \\ n\left(\frac{1-x}{1-\theta}\right)^{n-1}, & \text{for } \theta \leq x \leq 1. \end{cases}$$

- **Likelihood function for order statistics** $(X_{(1)}, \ldots, X_{(s)})$ is by definition:

$$L(\underline{X}\,;\theta\,,n) = n^s\{H(\underline{X};\theta)\}^{n-1}, \text{ where}$$

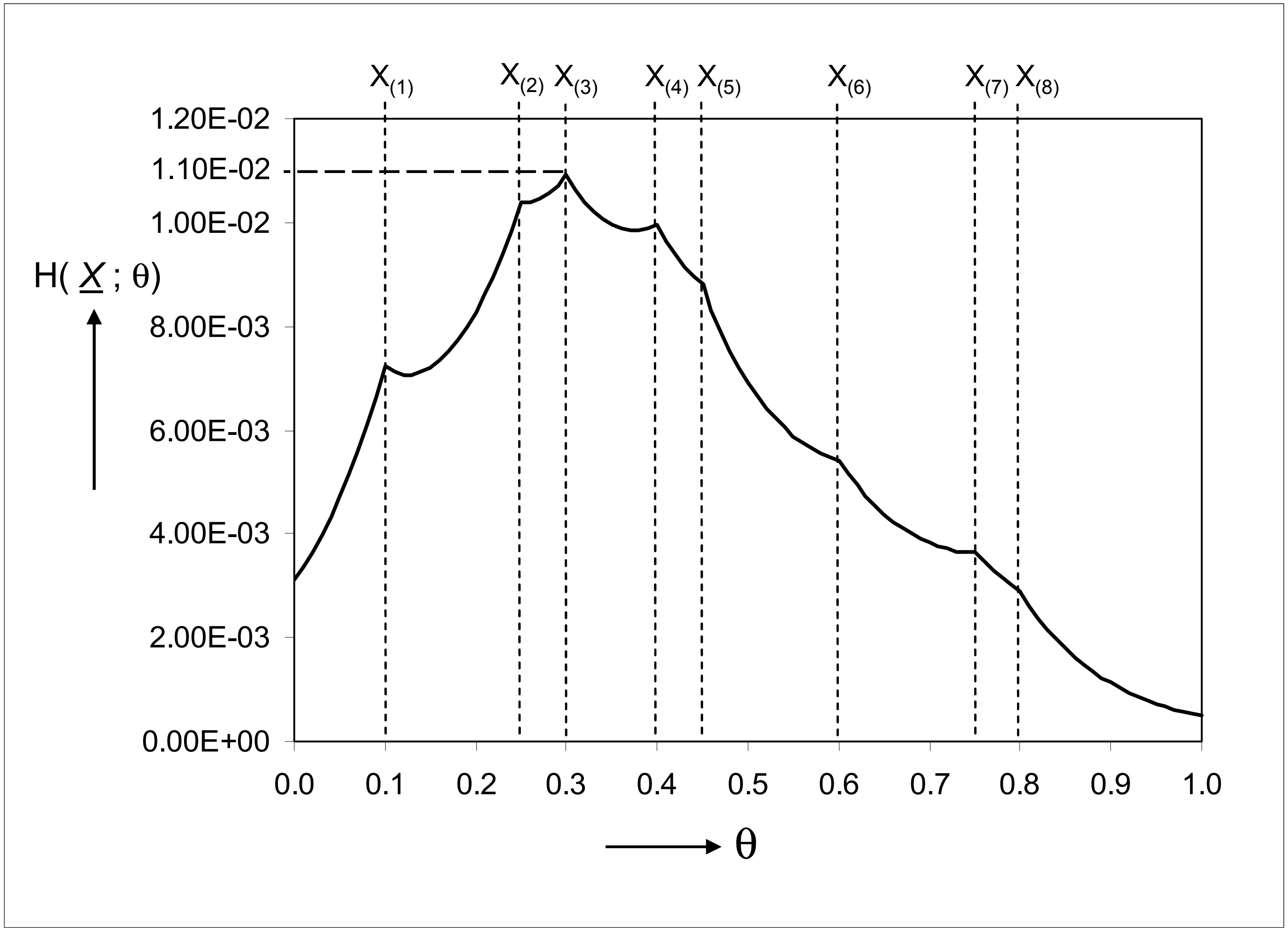
$$H(\underline{X};\theta) = \frac{\prod_{i=1}^{r} X_{(i)} \prod_{i=r+1}^{s} (1 - X_{(i)})}{\theta^r(1-\theta)^{s-r}}$$

and $X_{(r)} \leq \theta \leq X_{(r+1)}$ with $X_{(0)} \equiv 0,\ X_{(s+1)} \equiv 1.$



Graph of the likelihood $L(\underline{X};\theta,n)$ for the data

$$(X_{(1)},\ldots,X_{(8)}) = (0.10, 0.25, 0.30, 0.40, 0.45, 0.60, 0.75, 0.80)$$



Graph of $H(\underline{X}\,;\theta)$ for the same data (Maxima over the sets $[X_{(r)}, X_{(r+1)}]$ are attained here solely at the order statistics).

Theorem: *Let* $\underline{X} = (X_1, \ldots, X_s)$ *be an i.i.d. sample from a* $STSP(\theta, n)$ *distribution. The ML estimators of* θ *and* n *maximizing the likelihood* $L(\underline{X}\,; \theta\,, n)$ *over the parameter domain* $0 \leq \theta \leq 1$ *and* $n > 1$ *are:*

$$\begin{cases} \widehat{\theta} = X_{(\widehat{r})} \\ \widehat{n} = Max\left\{ -\frac{s}{Log\,M(\widehat{r})}, 1 \right\}, \end{cases}$$

$$\widehat{r} = \underset{r \in \{1, \ldots, s\}}{arg\,max} M(r)$$

$$M(r) = \prod_{i=1}^{r-1} \frac{X_{(i)}}{X_{(r)}} \prod_{i=r+1}^{s} \frac{1 - X_{(i)}}{1 - X_{(r)}}.$$

- **Modifications have to be made** when maximizing the likelihood $L(\underline{X}; \theta, n)$ over the parameter domain $0 \leq \theta \leq 1$ and $\mathbf{0 < n \leq 1}$.

30-YEAR US MORTGAGE INTEREST RATES

- Let the **interest rate after year *k*** be denoted by i_k; **one of the simplest financial engineering models** for the random behavior of the interest rates is the multiplicative model:

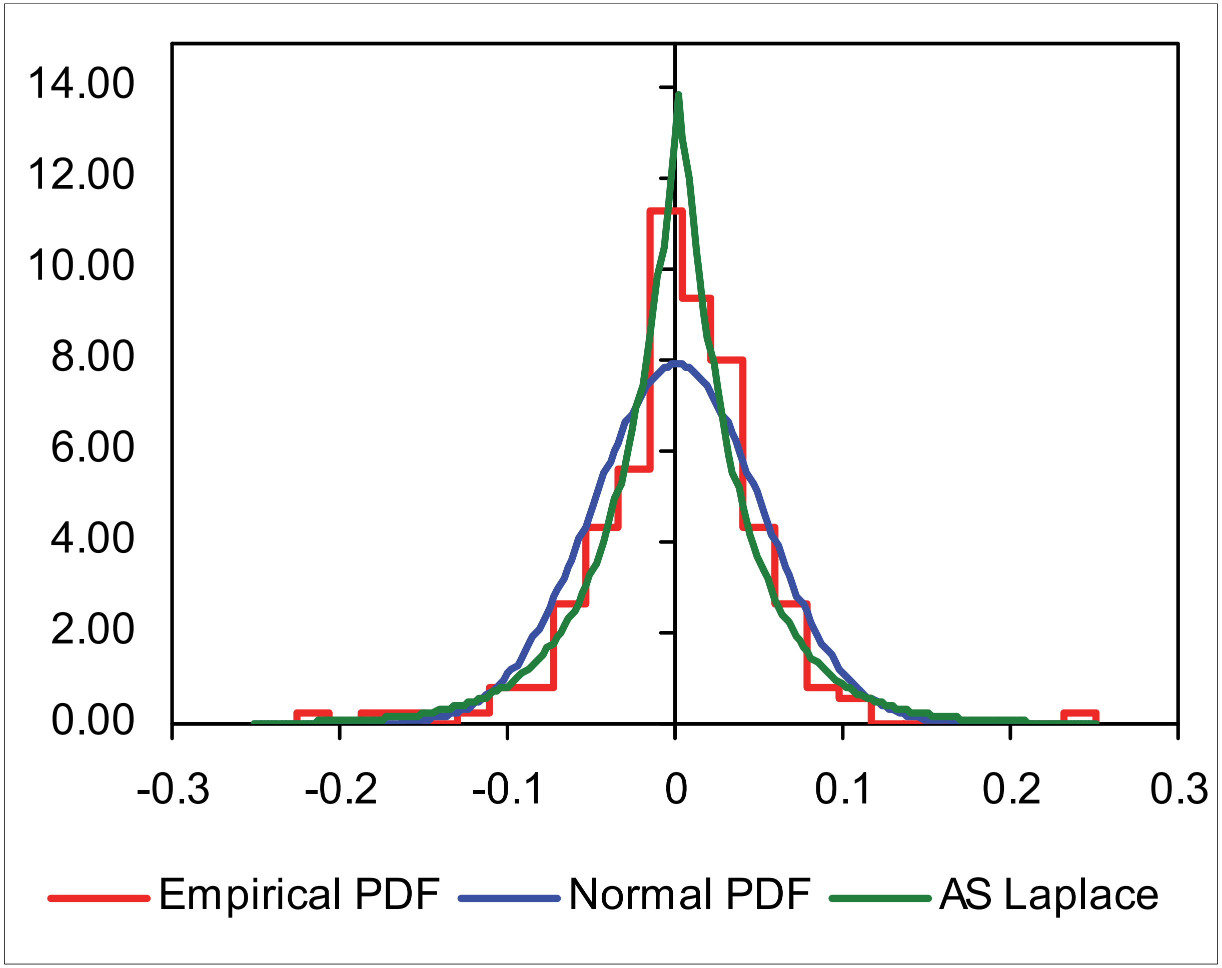
$$i_{k+1} = i_k \, \epsilon_k,$$

where ϵ_k are **identical independent distributed (i.i.d.)** random variables (see, e.g., Leunberger, D.G. (1998). *Investment Science.* New York, NY: Oxford University Press).

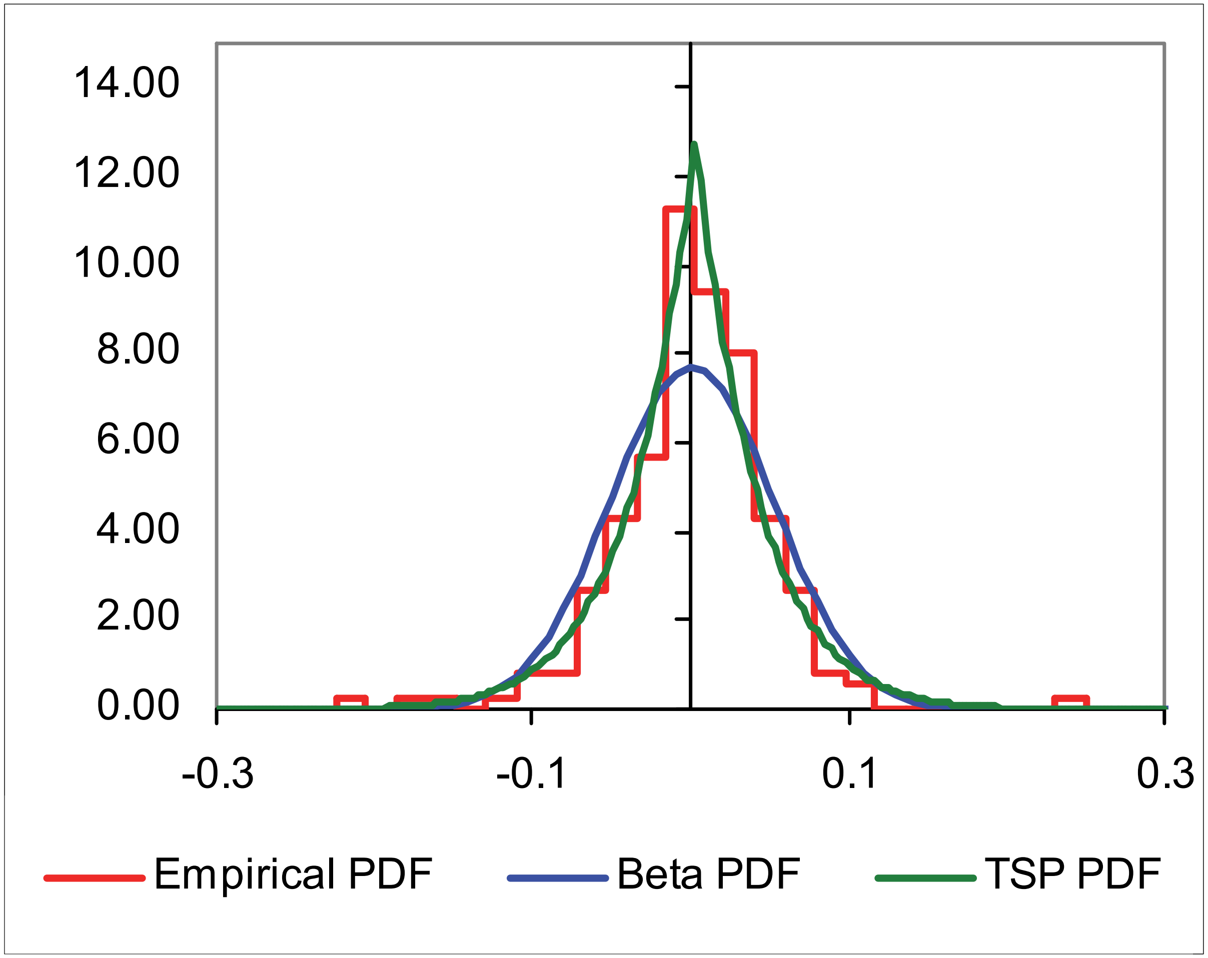
- Taking logarithms on LHS and RHS of yields:

$$Ln(i_{k+1}) = Ln(i_k) + Ln(\epsilon_k) \Leftrightarrow Ln(i_{k+1}) - Ln(i_k) = Ln(\epsilon_k)\,,$$

where as before the $Ln(\epsilon_k)$ are i.i.d.



Empirical pdf of **two-month lag log differences of 30-year conventional mortgage interest rates** together with the **ML fitted Gaussian** and **asymmetric Laplace distributions.**



Empirical pdf of **two-month lag log differences of 30-year conventional mortgage interest rates** together with **ML fitted beta** and **TSP distributions**.

SUMMARY OF MORE FORMAL FIT ANALYSIS

	Normal	AS Laplace	Beta	TSP
Chi-Squared Statistic	18.99	13.27	20.56	12.01
Degrees of Freedom	12	11	12	12
P-value	0.089	0.276	0.057	0.445
K-S Statistic	8.90%	6.44%	9.21%	4.66%
SS	0.276	0.091	0.320	0.062
Log-Likelihood	307.42	319.16	304.71	318.36

Except for the **Log-Likelihood Statistic** the
TSP distribution outperforms the other four distributions.

OUTLINE

1. Review properties of Triangular Distribution.
2. Early extensions of the Triangular Distribution.
3. The Two-Sided Power Distribution (TSP).
4. A link between TSP and Asymmetric Laplace distributions
5. Moment Ratio Diagram Comparison
6. Mazimum Likelihood Estimation STSP distributions
7. **A general framework of Two-Sided Distributions**
8. Generalized Trapezoidal Distributions
9. Generalized Two-Sided Power Distributions

- **STSP distributions** can be viewed as a particular case of the *general* Standard Two-Sided (STS) continuous family with support $[0,1]$ given by the density

$$g\{x|\theta, p(\,\cdot\,|\Psi)\} = \begin{cases} p(\frac{x}{\theta}|\Psi), & \text{for } 0 < x \leq \theta \\ p(\frac{1-x}{1-\theta}|\Psi), & \text{for } \theta < x < 1, \end{cases}$$

  where $\boldsymbol{p(\,\cdot\,|\Psi)}$ is an appropriately selected **continuous pdf defined on $[0,1]$ with parameter(s) Ψ**, which may in principle be vector-valued.

- The density $\boldsymbol{p(\,\cdot\,|\Psi)}$ will be referred to as the ***generating density*** of the resulting STS family of distributions and **the parameter θ** is termed the **reflection parameter** (an alternative designation could be the *hinge* or *threshold* parameter).

- **Triangular distribution** has generating density:

$$p(y) = 2y, \quad 0 \leq y \leq 1,$$

- **Standard Two-Sider Power distribution** has generating density:

$$p(y|n) = ny^{n-1}, \; 0 \leq y \leq 1, n > 0,$$

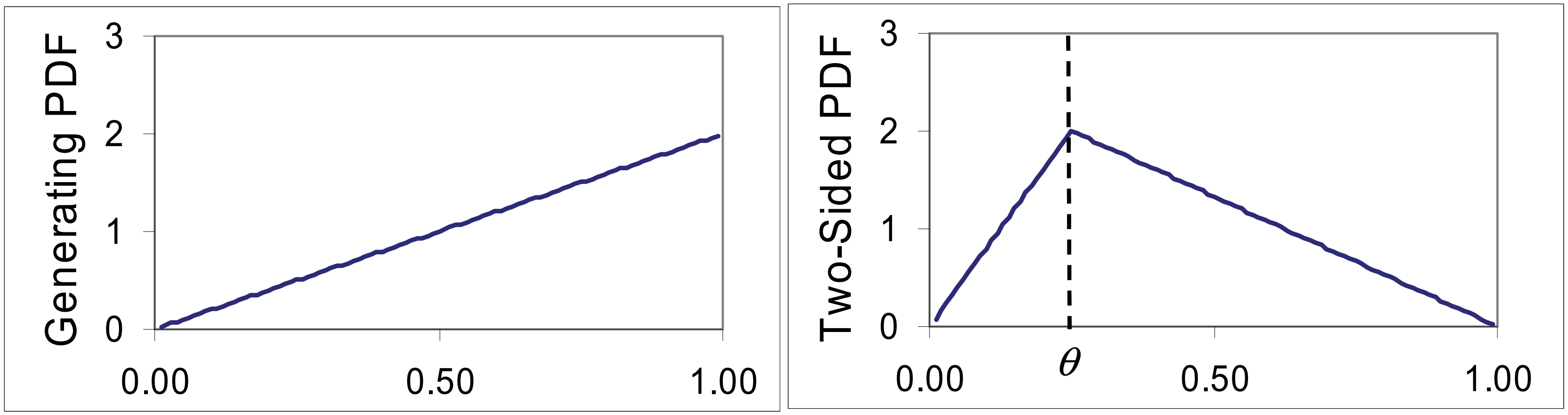
- **Two-Sided Slope distribution** has generating density:

$$p(y|\alpha) = \; \alpha + 2(1-\alpha)y, \; 0 \leq y \leq 1, \; 0 \leq \alpha \leq 2,$$

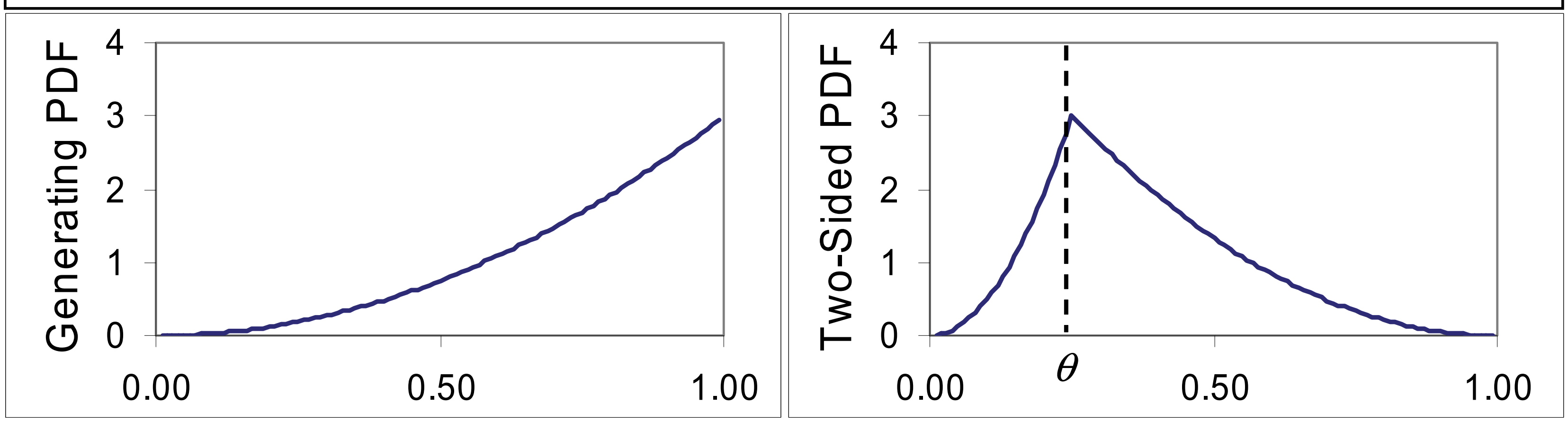
- **Two-Sided Ogive distribution** has generating density:

$$p(y|n) = ny^{n-1}\{\frac{4n-2}{3n-1} - \frac{2n-2}{3n-1}y^n\}, \; 0 \leq y \leq 1, n \geq 1/2.$$

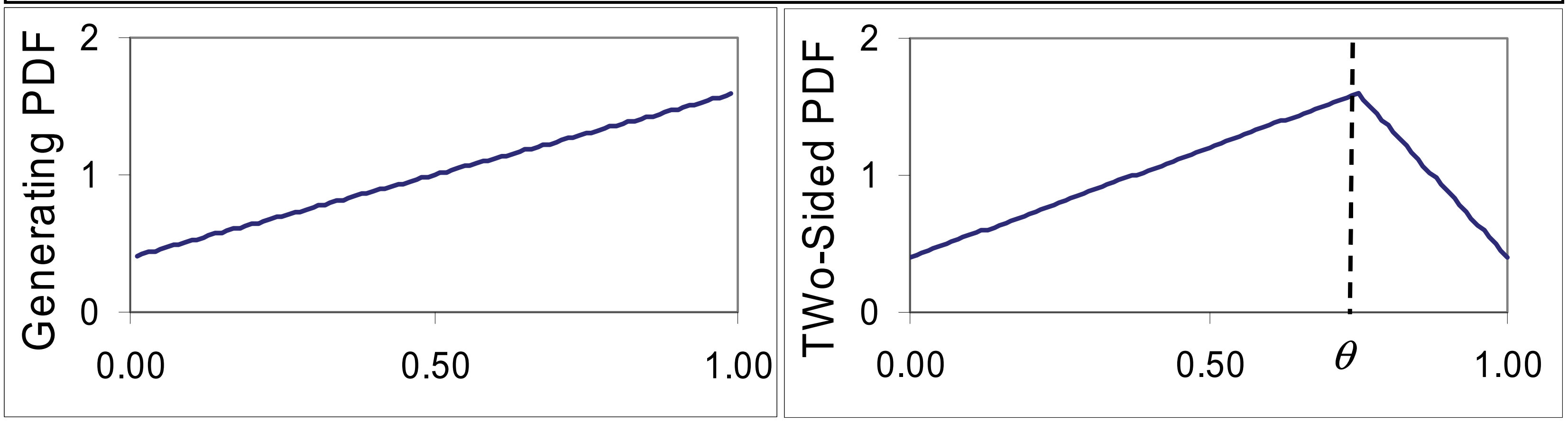
Triangular Distribution



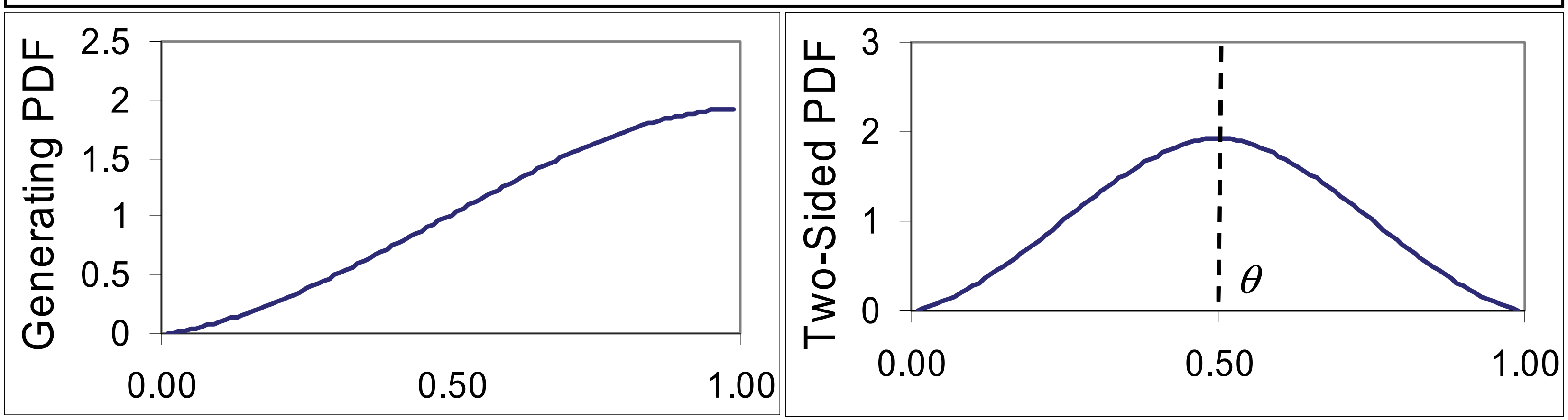
Two-Sided Power Distribution



Two-Sided Slope Distribution



Two-Sided Ogive Distribution



- $X \sim g\{\cdot|\theta, p(\cdot|\Psi)\}$ **two-sided distributed** with **generating density** $p(\cdot|\Psi)$ one obtains for **the cdf of** X

$$G\{x|\theta, P(\cdot\Psi)\} = \begin{cases} \theta P(\frac{x}{\theta}|\Psi), & \text{for } 0 < x < \theta \\ 1-(1-\theta)P(\frac{1-x}{1-\theta}|\Psi), & \text{for } \theta \leq x < 1, \end{cases}$$

where $P(\cdot|\Psi)$ is the cdf of the generating density $p(\cdot|\Psi)$.

- **An important and revealing property of the STS family** is that

$$G(\theta|\theta,\Psi) = \theta P(1|\Psi) = \theta.$$

- If $Y \sim p(\cdot|\Psi)$ with **moments** $\boldsymbol{E[Y^k]}$, $k = 1, 2, 3, \ldots$

$$E[X^k|\theta,\Psi] = \theta^{k+1}\boldsymbol{E[Y^k|\Psi]} + \sum_{i=0}^{k}\binom{k}{i}(-1)^i(1-\theta)^{i+1}\boldsymbol{E[Y^i|\Psi]}.$$

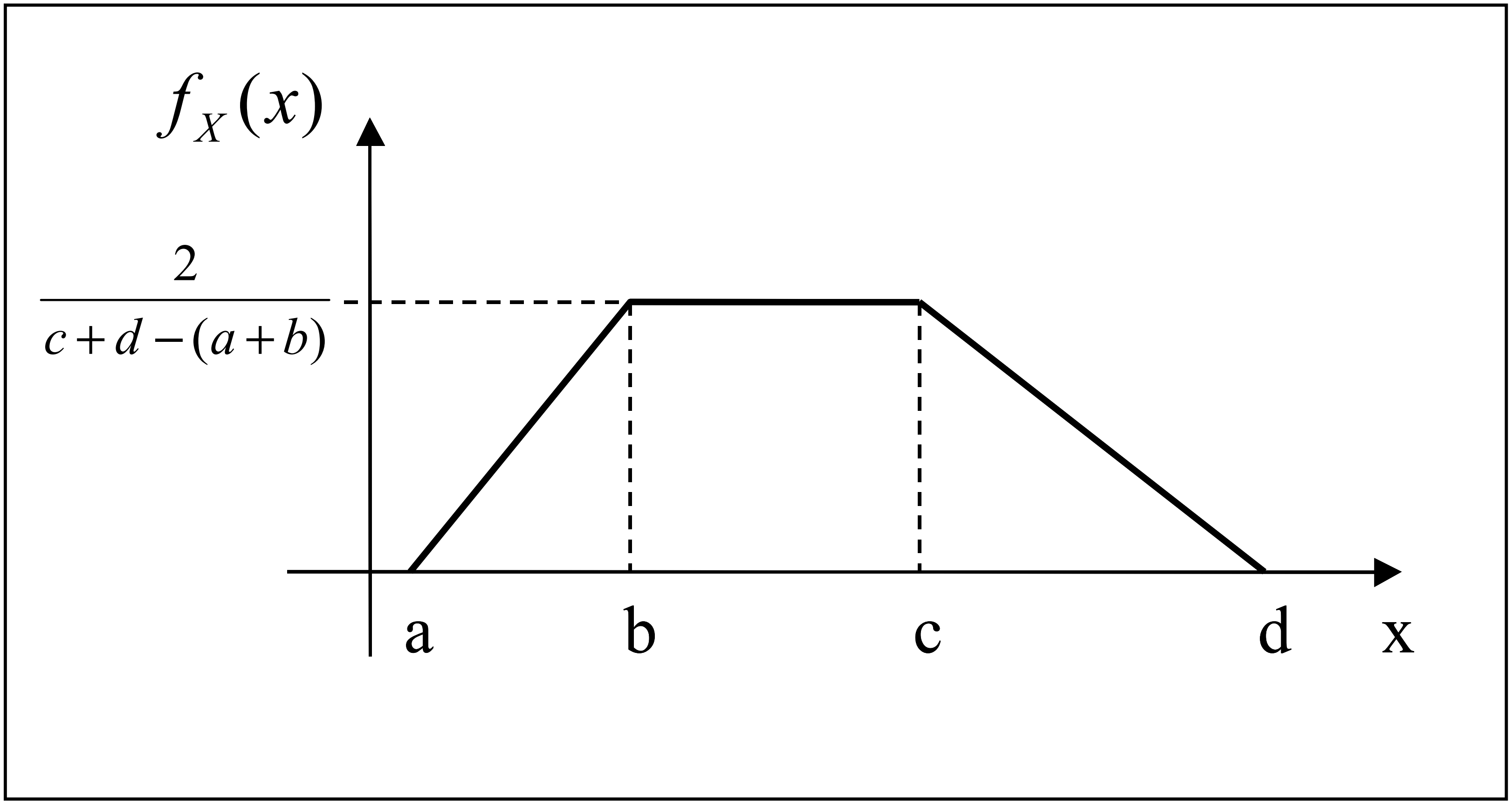
$$E[X|\theta,\Psi] = (2\theta-1)\boldsymbol{E[Y|\Psi]} + (1-\theta)$$

$$Var(X|\theta,\Psi) = \{1-3\theta+3\theta^2\}\boldsymbol{Var(Y|\Psi)} + \theta(1-\theta)\{\boldsymbol{E[Y|\Psi]}-1\}^2.$$

OUTLINE

1. Review properties of Triangular Distribution.
2. Early extensions of the Triangular Distribution.
3. The Two-Sided Power Distribution (TSP).
4. A link between TSP and Asymmetric Laplace distributions
5. Moment Ratio Diagram Comparison
6. Mazimum Likelihood Estimation STSP distributions
7. A general framework of Two-Sided Distributions
8. **Generalized Trapezoidal Distributions**
9. Generalized Two-Sided Power Distributions

Having generalized the Triangular distribution **can we generalize the Trapezoidal distribution** in a similar manner?



Probability density function of **a Trapezoidal Distribution.**

- Pdf of Generalized Trapezoidal Distributions:

$$f_X(x|a,b,c,d,m,n,\alpha) =$$
$$\begin{cases} \frac{2\alpha mn}{2\alpha(b-a)n+(\alpha+1)(c-b)mn+2(d-c)m}\left(\frac{x-a}{b-a}\right)^{m-1} & a \le x < b \\ \frac{2mn}{2\alpha(b-a)n+(\alpha+1)(c-b)mn+2(d-c)m}\left\{(\alpha-1)\frac{c-x}{c-b}+1\right\} & b \le x < c \\ \frac{2mn}{2\alpha(b-a)n+(\alpha+1)(c-b)mn+2(d-c)m}\left(\frac{d-x}{d-c}\right)^{n-1} & c \le x < d \end{cases}$$

where $m > 0, n > 0,\ \alpha > 0$ and $a < b < c < d.$

- The pdf above is constructed **using three probabilities** $\sum_{i=1}^{3}\pi_i = 1,$ $\pi_i > 0$ and **a mixture of three densities**

$$f_X(x) = \sum_{i=1}^{3}\pi_i f_{X_i}(x).$$

- Setting **the mixture density elements** equal to

$$f_{X_1}(x|a,b,m) = \left(\frac{m}{b-a}\right)\left(\frac{x-a}{b-a}\right)^{m-1}, a \leq x < b, m > 0,$$

$$f_{X_2}(x|b,c,\alpha) = \frac{2}{(\alpha+1)(c-b)^2}\{(1-\alpha)x + \alpha c - b\}, b \leq x \leq c, \alpha > 0,$$

$$f_{X_3}(x|c,d,n) = \left(\frac{n}{d-c}\right)\left(\frac{d-x}{d-c}\right)^{n-1}, c \leq x < d, n > 0$$

the problem of solving for the mixture probabilities, $\pi_1, \pi_2, \pi_3 > 0$, such that **the density $f_X(x)$ is continuous over the support $[a, d]$** is not trivial. See:

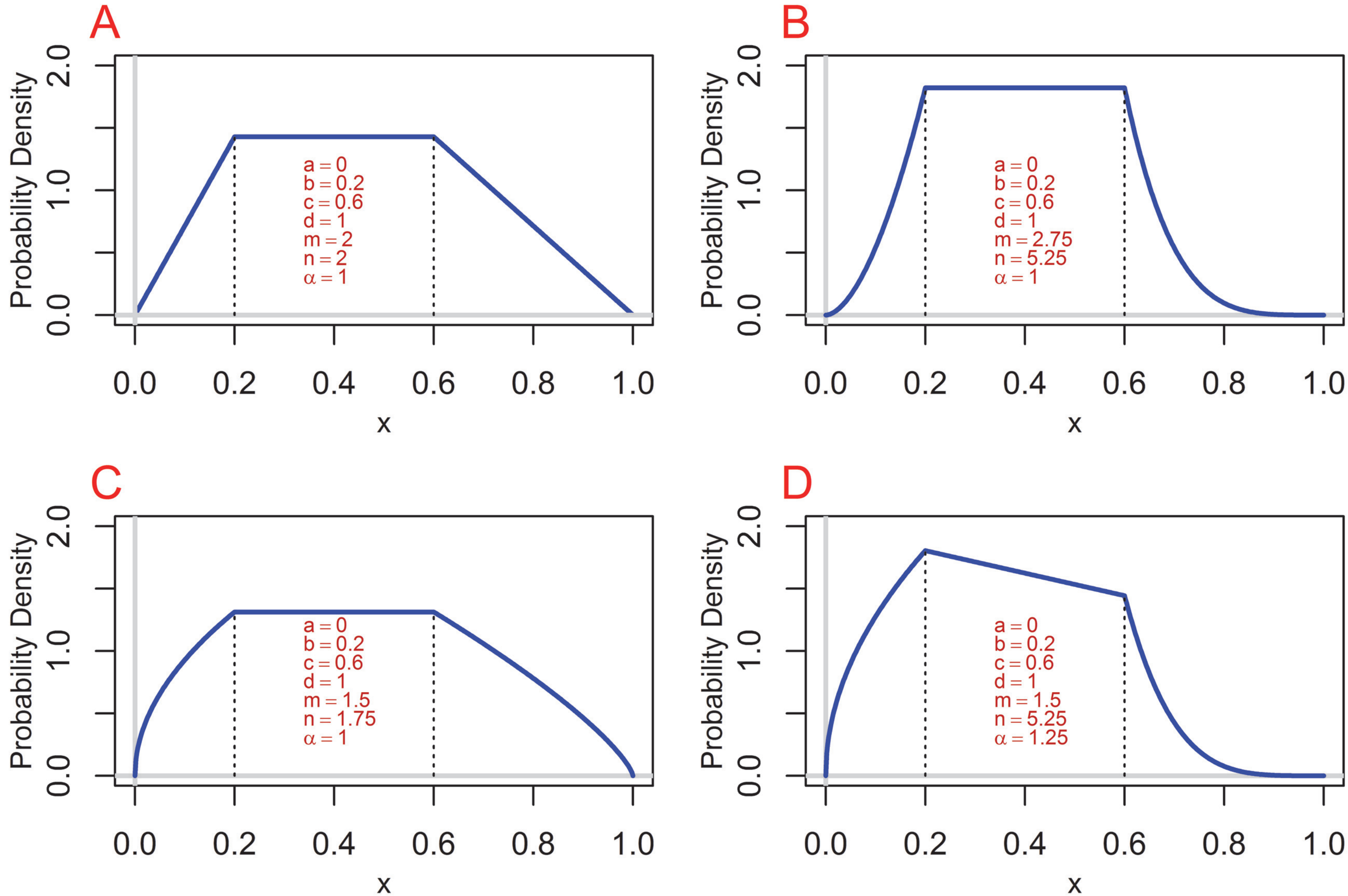
J.R. van Dorp and S. Kotz (2003). "Generalized Trapezoidal Distributions". *Metrika*, Vol. 58, Issue 1, pp. 85-97. **R-package for Generalized Trapezoidal distribution is available on my faculty page**, Courtesy of Jeremy Thoms Hetzel (April 25, 2011).

Proposition: *The generelized trapezoidal probability density function follows from expressions for* $f_{X_1}(x|a,b,m)$, $f_{X_2}(x|b,c,\alpha)$, $f_{X_3}(x|c,d,n)$ *utilizing mixture probabilities*

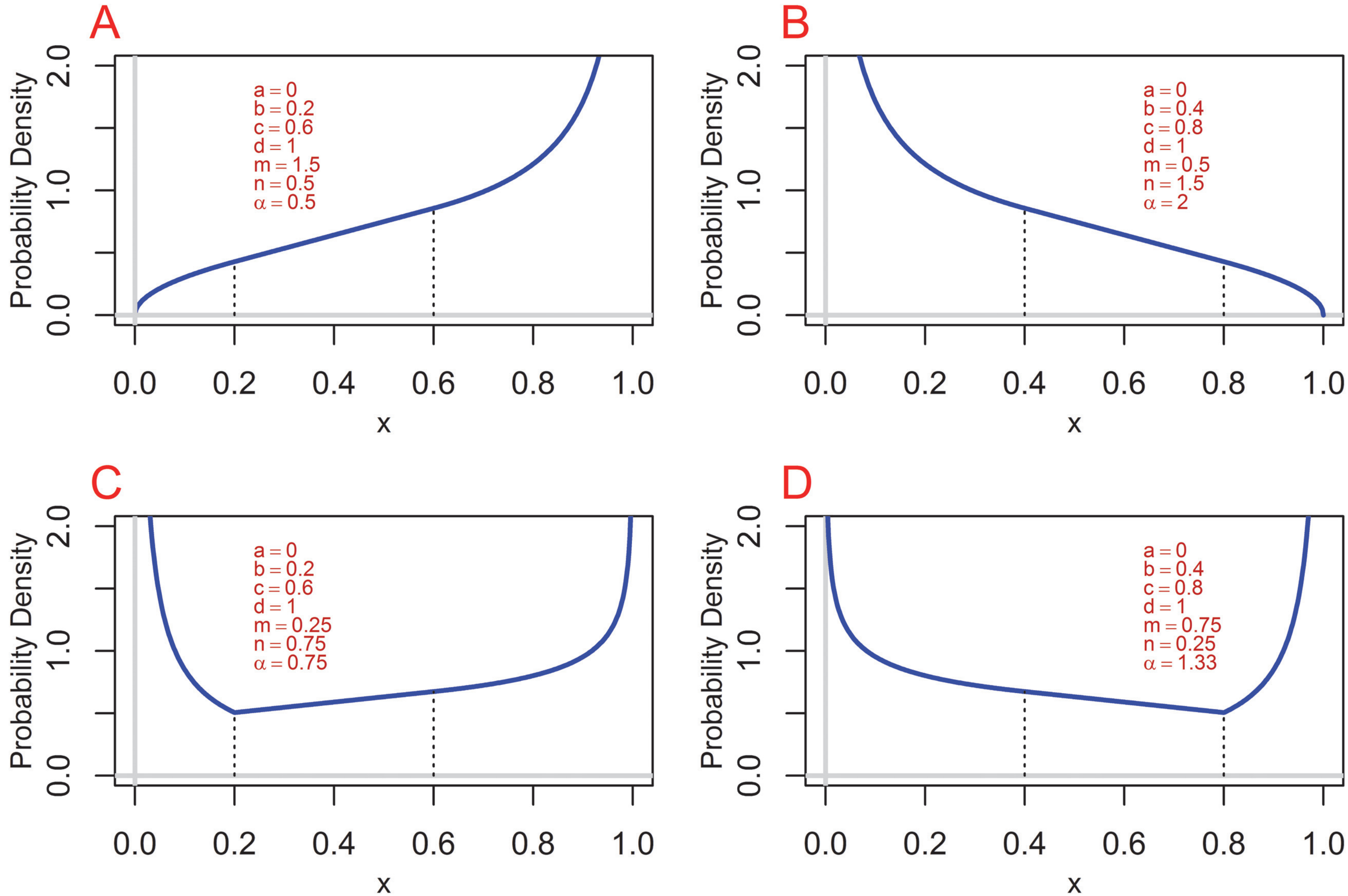
$$\begin{cases} \pi_1 = \frac{2\alpha(b-a)n}{2\alpha(b-a)n+(\alpha+1)(c-b)mn+2(d-c)m} \\ \pi_2 = \frac{(\alpha+1)(c-b)mn}{2\alpha(b-a)n+(\alpha+1)(c-b)mn+2(d-c)m} \\ \pi_3 = \frac{2(d-c)m}{2\alpha(b-a)n+(\alpha+1)(c-b)mn+2(d-c)m}, \end{cases}$$

where $a < b < c < d,\ n_1 > 0, n_2 > 0, \alpha > 0$ *and the probability density function given by* (3) *is continuous.*

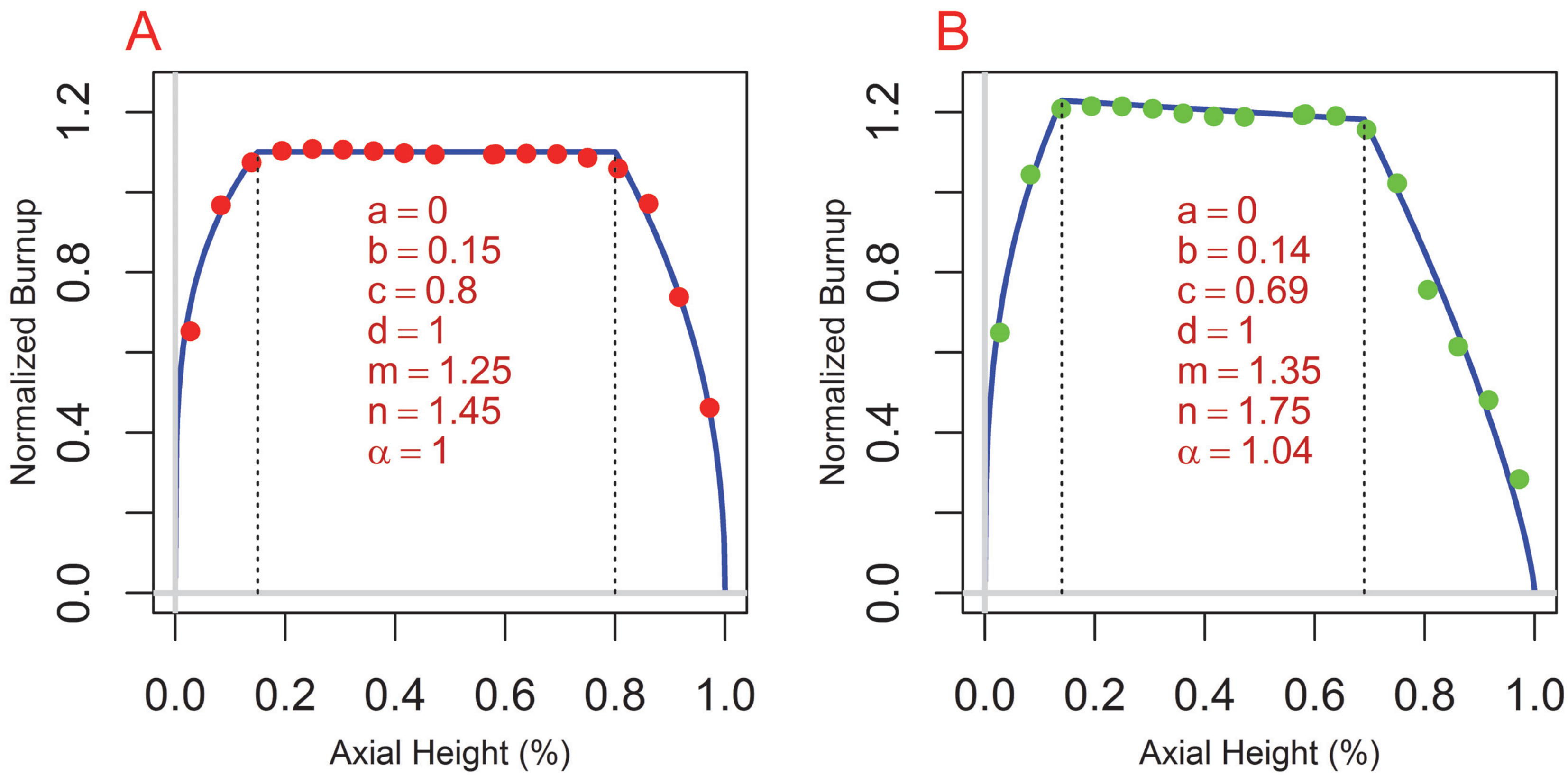
- Note that the numerator of **the first stage probability π_1** has as a multiplier **the power parameter of the third stage**.
- Note that the numerator of **the third stage probability π_3** has as a multiplier **the power parameter of the first stage.**
- Note that the numerator of **the second stage probability π_2** has both power parameters of **the first and third stage as multipliers.**


A
a = 0
b = 0.2
c = 0.6
d = 1
m = 2
n = 2
α = 1
B
a = 0
b = 0.2
c = 0.6
d = 1
m = 2.75
n = 5.25
α = 1
C
a = 0
b = 0.2
c = 0.6
d = 1
m = 1.5
n = 1.75
α = 1
D
a = 0
b = 0.2
c = 0.6
d = 1
m = 1.5
n = 5.25
α = 1.25
Probability Density
x
0.0
0.2
0.4
0.6
0.8
1.0
1.0
2.0

J-Shaped and U-Shaped



While fun mathematical exercises, these new distribution models are **only usefull** if they can find application **better than other models.**



Data From: Wagner JC, DeHart MD (2000). Review of Axial Burnup Distributions for Burnup Credit Calculations. *Oak Ridge National Laboratory*, ORNL/TM-1999/246, Oak Ridge, Tennessee.

OUTLINE

1. Review properties of Triangular Distribution.
2. Early extensions of the Triangular Distribution.
3. The Two-Sided Power Distribution (TSP).
4. A link between TSP and Asymmetric Laplace distributions
5. Moment Ratio Diagram Comparison
6. Mazimum Likelihood Estimation STSP distributions
7. A general framework of Two-Sided Distributions
8. Generalized Trapezoidal Distributions
9. **Generalized Two-Sided Power Distributions**

- Setting $a = 0, b = c = \theta, \alpha = 1$ in the Generalized Trapezoidal distribution, we arrrive at the pdf of a **Generalized Two-Sided Power (GTSP) distribution** that allows for **seperate power parameters within each of its branches** $n, m > 0, 0 \leq \theta \leq 1$:

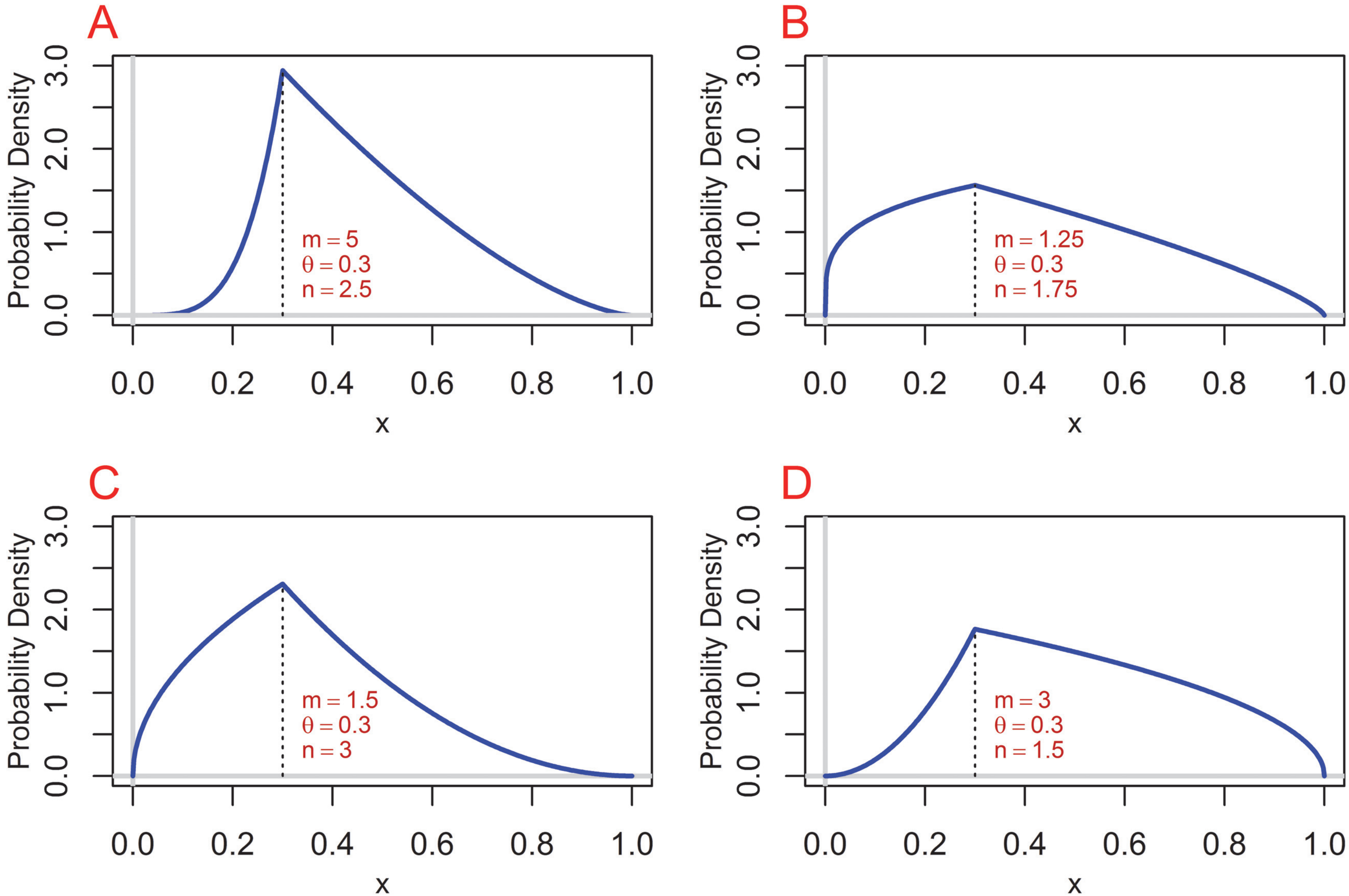
$$f_X(x|\Theta) = \frac{mn}{(1-\theta)m + \theta n} \times \begin{cases} \left(\frac{x}{\theta}\right)^{m-1}, & \text{for } 0 < x < \theta \\ \left(\frac{1-x}{1-\theta}\right)^{n-1}, & \text{for } \theta \leq x < 1. \end{cases}$$

- **The mode probability** is now also a function of the power parameters

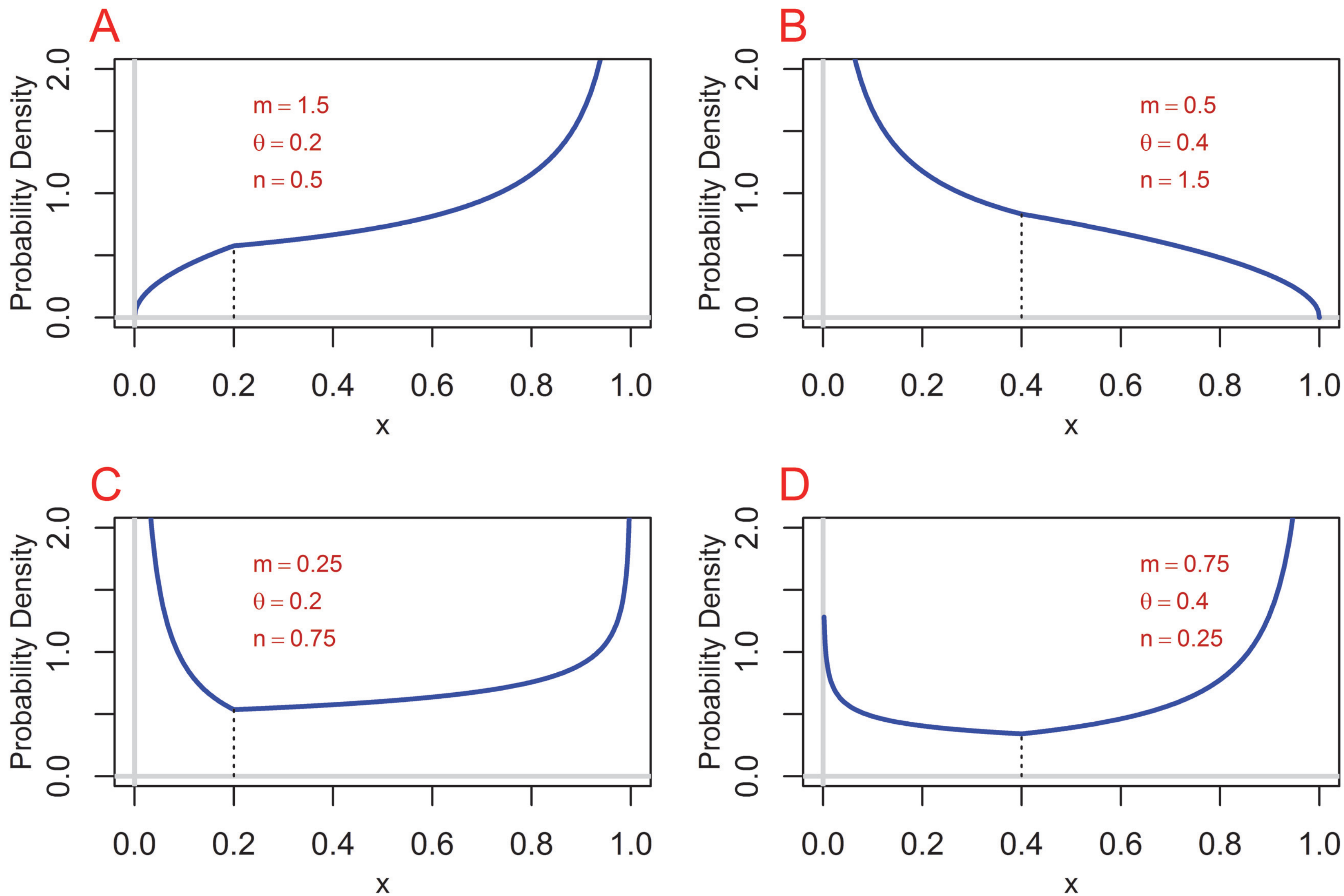
$$Pr(X \leq \theta) = \frac{\theta n}{(1-\theta)m + \theta n}$$

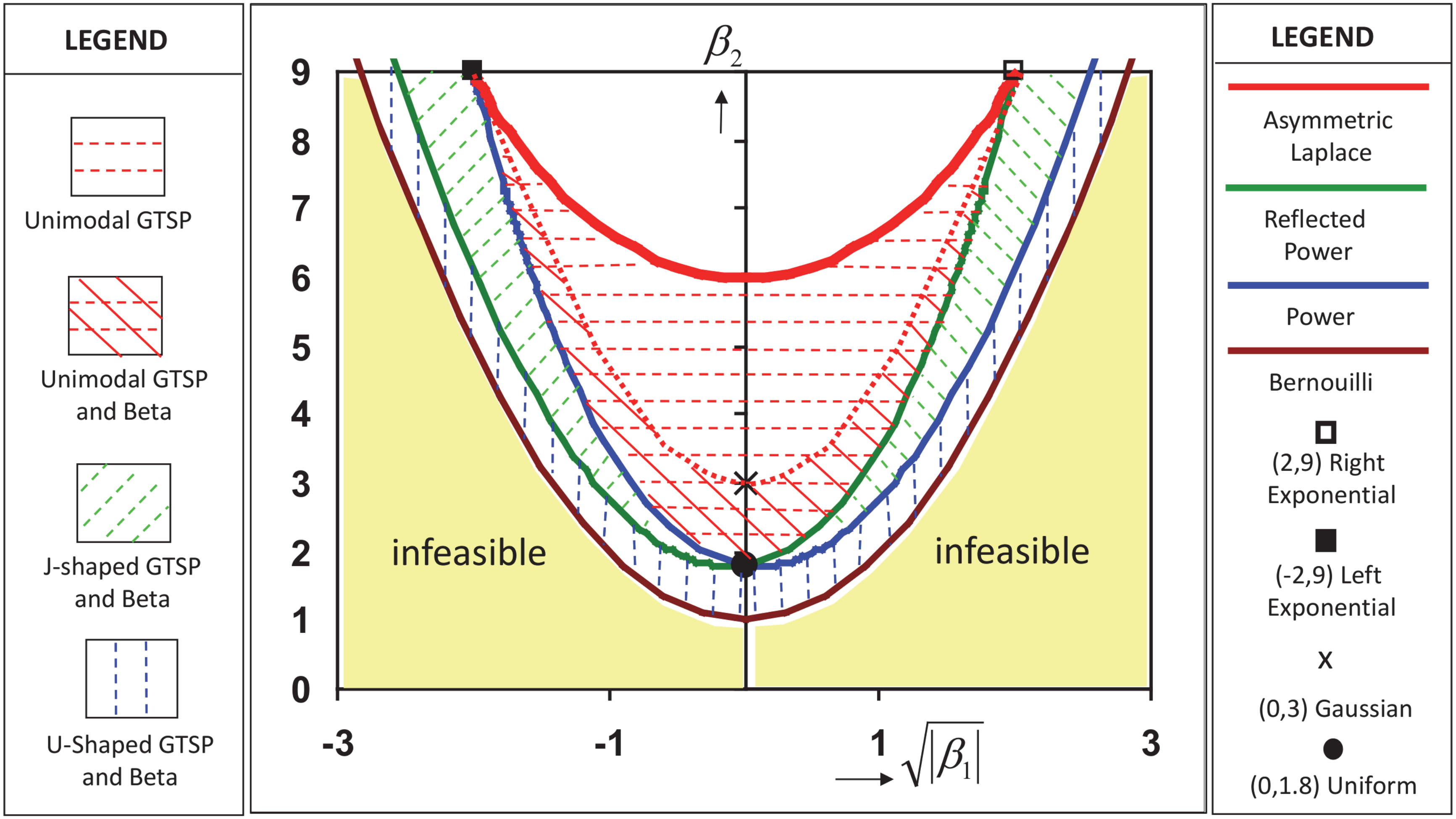
- **Mean, Variance and other moment expressions** are more involved.

$$E[X] = m\frac{n(n+1)\theta^2 + (m+1)(1-\theta)(n\theta + 1)}{(m+1)(n+1)\{(1-\theta)m + \theta n\}}$$


A
Probability Density
x
m = 5
θ = 0.3
n = 2.5
B
m = 1.25
θ = 0.3
n = 1.75
C
m = 1.5
θ = 0.3
n = 3
D
m = 3
θ = 0.3
n = 1.5
0.0 1.0 2.0 3.0
0.0 0.2 0.4 0.6 0.8 1.0

8. GTSP DISTRIBUTIONS... Unimodal Shapes





GTSP Distribution now **shares the J-Shaped coverage with the Beta distribution** that the **TSP distribution did not cover.**

BEYOND BETA

Statistical distributions are fundamental to statistical science and are a prime indispensable tool for its applications. The BEYOND BETA monograph is the first to examine an important but somewhat neglected area in this field - univariate continuous distributions on a bounded domain, excluding the beta distribution. It provides an elementary but thorough discussion of "novel" contributions developed in recent years and some of their applications. It contains a comprehensive chapter on the triangular distribution as well as a chapter on earlier extensions of this distribution not emphasized in existing statistical literature. Special attention is given to estimation, in particular, non-standard maximum likelihood procedures. In the presentation, we shall concentrate on the properties and parameter estimation of the Two-Sided Power (TSP) distributions and their link to Generalized Trapezoidal and the Asymmetric Laplace distributions. The different shapes of TSP pdf's are presented on the book cover and mimic those of the beta distributions. The general framework for Two-Sided distribution will be presented as well.

Publication Date: December 15, 2004

Co-authored by **Samuel Kotz** and **Johan Rene van Dorp**, both at **The George Washington University**, Department of Engineering Management and Systems Engineering, Washington, D.C., USA

Samuel Kotz is the co-author with Dr. N.L. Johnson of the recently published *Leading Personalities in Statistical Sciences,* the classic five-volume *Distributions in Statistics* (now in its second edition), amongst other books in Statistics and Probability.

J. Rene van Dorp is active in applications of statistical methodology such as probabilistic risk analysis, reliability and simulation, and is an author of over 20 scientific publications.

World Scientific
Connecting Great Minds
http://www.worldscientific.com/