Lecture 1

INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

Class: TY Bsc

Subject : Statistical Techniques and Risk Management 4

Subject Code: PUSASQF605A

Chapter: Unit 2

Chapter Name: Extreme Value Theory

Chapter Agenda

1. Introduction to Extreme Events
    1. Heteroscedasticity
    2. Block Maxima
    3. Peak Over Threshold

2. General Extreme Value Distribution

3. General Pareto Distribution

4. Measure of Tail Weights

1. Introduction to Extreme Events

Extreme Events

- Events that have a low probability of occurring and a high financial impact i.e. low frequency and high severity.

Example :-

> A loss due to natural calamity ; stock market crash.

> A large amount of claims come together

- As we extrapolate extreme value probabilities from our median data we end up understating them.

1. Introduction to Extreme Events

Extreme Events

- Problems with modelling extreme events

1. No data / very less data for extreme events.
2. Non normality of financial returns or losses as these event distributions have fatter tails and sharper peaks
3. Probability of extreme events is underestimated as normal distribution has narrow tails.

1. Introduction to Extreme Events

Extreme Events

- Kurtosis

K = 3 is mesokurtic distribution
[Normal Distribution]

K > 3 is leptokurtic distribution
[Sharper Peak]

K < 3 is platykurtic distribution
[Flatter Peak]

1.1 Heteroscedasticity

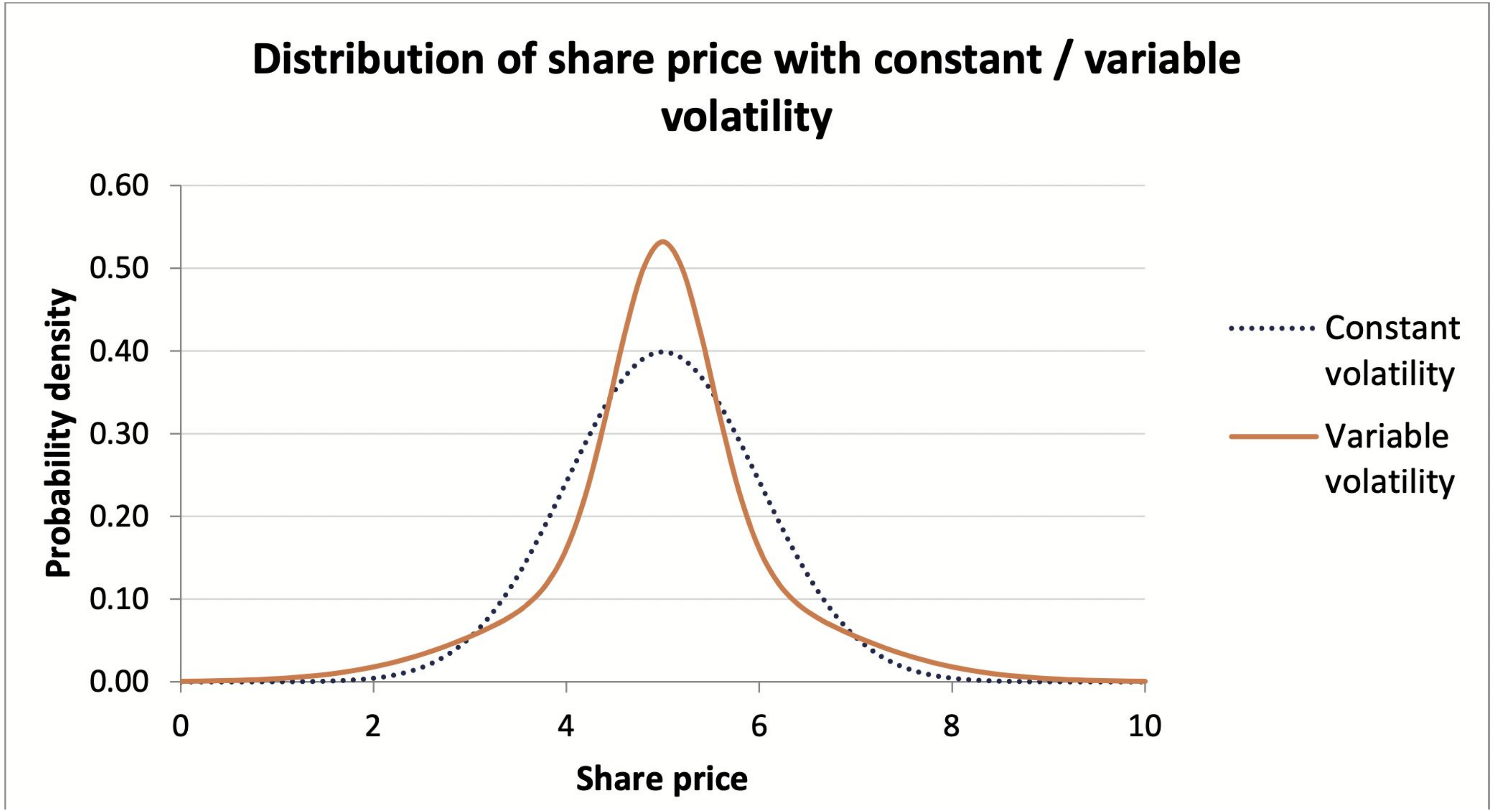
Asset Return Volatility is not a constant (clustering volatility) but it changes stochastically with time

- Heteroscedastic pdf has a fatter tail than pdf with constant volatility

The graph below compares two distributions for the price of a share in one year's time:

- a $N(5, \sigma^2)$ distribution with constant volatility, $\sigma = 1$
- a $N(5, \sigma^2)$ distribution where the volatility is heteroscedastic, *ie* $\sigma = 0.5$ and $\sigma = 1.5$ with equal probability.

1.1 Heteroscedasticity



1.1 Heteroscedasticity

Extreme Values

- Block Maxima

  Divide the data into blocks and select the highest value from each block.

- Peak Over Threshold Exceedances

  Model the data above a threshold level

1.2 Block Maxima

The dataset below shows the claim amounts in £000s in respect of a commercial property portfolio over a period of a year.

Claim number	*Claim amount*
1	9
2	28
3	20
4	8
5	102
6	152
7	23
8	108
9	42
10	12
11	110
12	9
13	22
14	37
15	147
16	128

Claim number	*Claim amount*
17	12
18	35
19	12
20	75
21	80
22	42
23	9
24	122
25	145
26	13
27	16
28	113
29	9
30	8
31	12
32	84

Claim number	*Claim amount*
33	19
34	17
35	66
36	55
37	81
38	140
39	64
40	9
41	9
42	36
43	185
44	135
45	25
46	16
47	55
48	31

Claim number	*Claim amount*
49	118
50	55
51	14
52	94
53	54
54	81
55	62
56	83
57	23
58	19
59	55
60	104

1.2 Block Maxima

(i) Determine the values of X_M where the block size is:

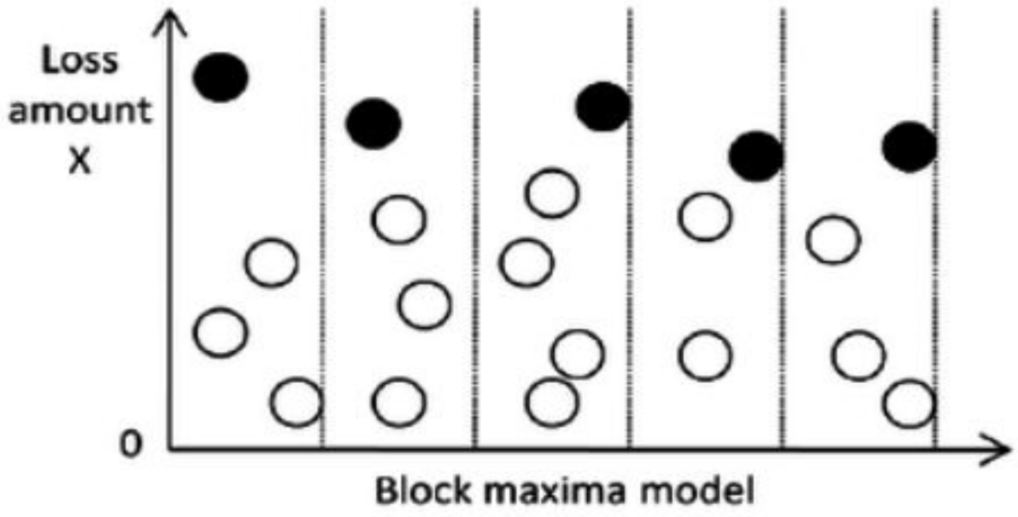
(a) $n = 5$

(b) $n = 10$

(ii) Comment on the trade-off between the block size and the values of X_M that will be used to fit the extreme value distribution.

1.2 Block Maxima

X_m = max value of X in a block
m = no of blocks
n = no of values in a block
$X_m = \max(X_1, X_2, \ldots X_n)$
X_m is the block maxima



1.2 Block Maxima

Generalized Extreme Value Distributions (GEV)

Whatever the underlying distribution of data the distribution of the standardized maximum values will converge to a distribution called General Extreme Value Distribution.

2. General Extreme Value Distribution

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \rightarrow \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

$$X_m = \max(X_1, X_2 \ldots X_n) \rightarrow \text{Block Maxima}$$

Model

$$\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)\ [standardised]$$

for extreme values

$$\frac{X_m - \alpha_n}{\beta_n} \rightarrow Modelling\ Upper\ Tail$$

2. General Extreme Value Distribution

CDF of X_m

2. General Extreme Value Distribution

X ~ Exp(λ)

$\alpha_n = \frac{1}{\lambda} \ln n$

$\beta_n = \frac{1}{\lambda}$

$\lim_{n\to\infty} P\left(\frac{X_m - \alpha_n}{\beta_n} \leq x\right)$

$\lim_{n\to\infty} P(X_m \leq x * \beta_n + \alpha_n)$

2. General Extreme Value Distribution

GEV distribution

$$\text{GEV} \rightarrow \begin{array}{lll} \alpha \ (location\ parameter\ ) & \rightarrow mean \\ \beta \ (scale\ parameter\ ) & \rightarrow variance \\ \gamma \ (shape\ parameter\ ) & \rightarrow skewness \end{array}$$

CDF of GEV distribution :

$$H(x) = \begin{cases} \exp\left(-\left(1+\dfrac{\gamma(x-\alpha)}{\beta}\right)^{-\frac{1}{\gamma}}\right) & \gamma \neq 0 \\ \exp\left(-\exp\left(-\dfrac{(x-\alpha)}{\beta}\right)\right) & \gamma = 0 \end{cases}$$

2. General Extreme Value Distribution

Gumbel Type GEV

Based on γ, there are 3 types of GEV Distribution

1. $\gamma = 0 \rightarrow Gumbel\ Type\ Distribution$
2. $\gamma > 0 \rightarrow Fretchet\ GEV$
3. $\gamma < 0 \rightarrow Weibull\ GEV$

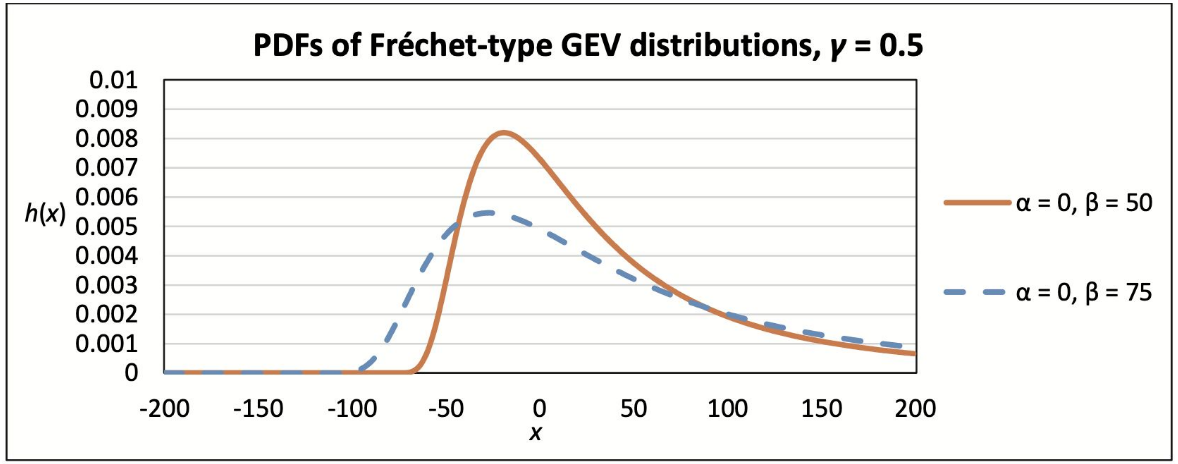
2. General Extreme Value Distribution

PDF of GEV distribution:

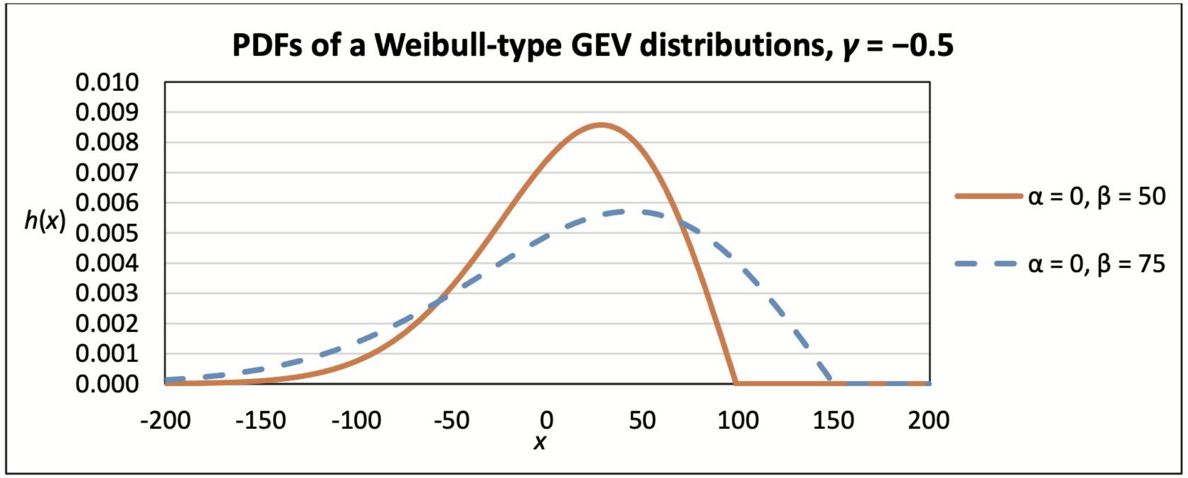
2. General Extreme Value Distribution

PDF of GEV distribution:

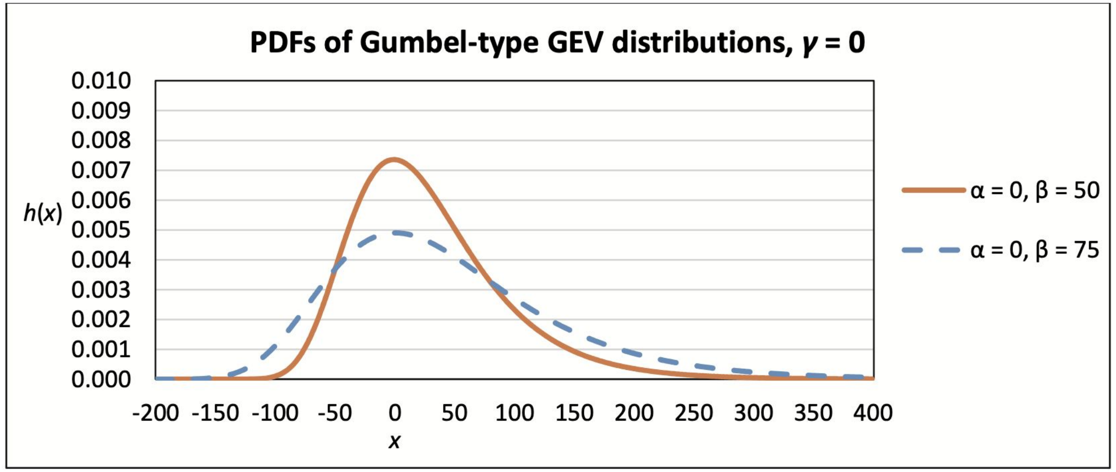
Fréchet-type GEV distributions



Weibull-type GEV distribution



Gumbel-type GEV distributions



2. General Extreme Value Distribution

$\gamma > 0 \ \rightarrow Fretchet$

- Highly +vely skewed

$\rightarrow heavy\ tailed$

Used for modelling extreme financial event lower bound X>$\left(\alpha - \frac{\beta}{\gamma}\right)$

$\gamma < 0$

- Highly −vely skewed

- Finite upper limit = upper bound $X < \alpha - \frac{\beta}{\gamma}$

$\gamma = 0$

- mildly +ve skewed
- no bounds

2. General Extreme Value Distribution

Choosing form of GEV distribution:

	GEV distributions (for the maximum value) corresponding to common loss distributions		
Type	**WEIBULL**	**GUMBEL**	**FRÉCHET**
Shape parameter	$\gamma < 0$	$\gamma = 0$	$\gamma > 0$
Underlying distribution	**Beta** **Uniform** **Triangular**	**Chi-square** **Exponential** **Gamma** **Lognormal** **Normal** **Weibull**	**Burr** ***F*** **Log-gamma*** **Pareto** ***t***
Range of values permitted	$x < \alpha - \frac{\beta}{\gamma}$	$-\infty < x < \infty$	$x > \alpha - \frac{\beta}{\gamma}$

2. General Extreme Value Distribution

The claim amounts in a general insurance portfolio are independent and follow an exponential distribution with mean £2,500.

(i) Calculate the probability that an individual claim will exceed £10,000. [1]

(ii) Calculate the probability that, in a sample of 100 claims, the largest claim will exceed £10,000 using:

(a) an exact method

(b) an approximation based on a Gumbel-type GEV distribution. [5]

You are given that, for an exponential distribution with parameter λ, the approximate distribution of $\max\{X_1,\ldots,X_n\}$ *for large n is a Gumbel-type GEV distribution with CDF:*

$$H(x)=\exp\left(-\exp\left(-\left(\frac{x-\alpha_n}{\beta_n}\right)\right)\right) \quad \textit{where } \alpha_n=\frac{1}{\lambda}\ln n \textit{ and } \beta_n=\frac{1}{\lambda}$$

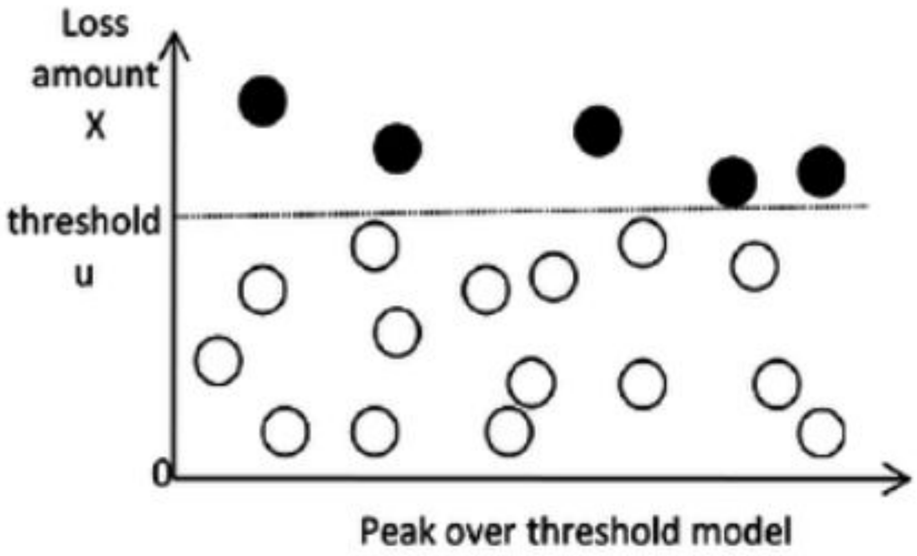
(iii) State the two key assumptions made in (ii)(a). [1]

[Total 7]

3 Peak Over Threshold Exceedances

Generalized Pareto Distribution

- Whatever the underlying distribution the distribution of the threshold exceedances will converge on the distribution called Generalized Pareto Distribution.



3. Generalised Pareto Distribution

Threshold Exceedances

$X - u \,/ X > u \qquad u \rightarrow threshold$

The higher the value of the threshold, the more extreme values of X. However using a higher threshold means that we have fewer values with which to fit the extreme value distribution.

3. Generalised Pareto Distribution

CDF of $X - u \,/ X > u$

3. Generalised Pareto Distribution

CDF of $X - u \,/ X > u$

3. Generalised Pareto Distribution

The dataset below shows the claim amounts in £000s in respect of a commercial property portfolio over a period of a year. (This is the dataset from the question in Section 2.1.)

Claim number	*Claim amount*	*Claim number*	*Claim amount*	*Claim number*	*Claim amount*	*Claim number*	*Claim amount*
1	9	17	12	33	19	49	118
2	28	18	35	34	17	50	55
3	20	19	12	35	66	51	14
4	8	20	75	36	55	52	94
5	102	21	80	37	81	53	54
6	152	22	42	38	140	54	81
7	23	23	9	39	64	55	62
8	108	24	122	40	9	56	83
9	42	25	145	41	9	57	23
10	12	26	13	42	36	58	19
11	110	27	16	43	185	59	55
12	9	28	113	44	135	60	104
13	22	29	9	45	25		
14	37	30	8	46	16		
15	147	31	12	47	55		
16	128	32	84	48	31		

(i) Calculate the values of $X-u\,|\,X>u$ when:

(a) $u=100$

(b) $u=125$

3. Generalised Pareto Distribution

With reference to the question above, use the method of maximum likelihood to fit a distribution to the threshold exceedances when the underlying claims distribution is exponential and the threshold is chosen to be 100.

3. Generalised Pareto Distribution

More generally,

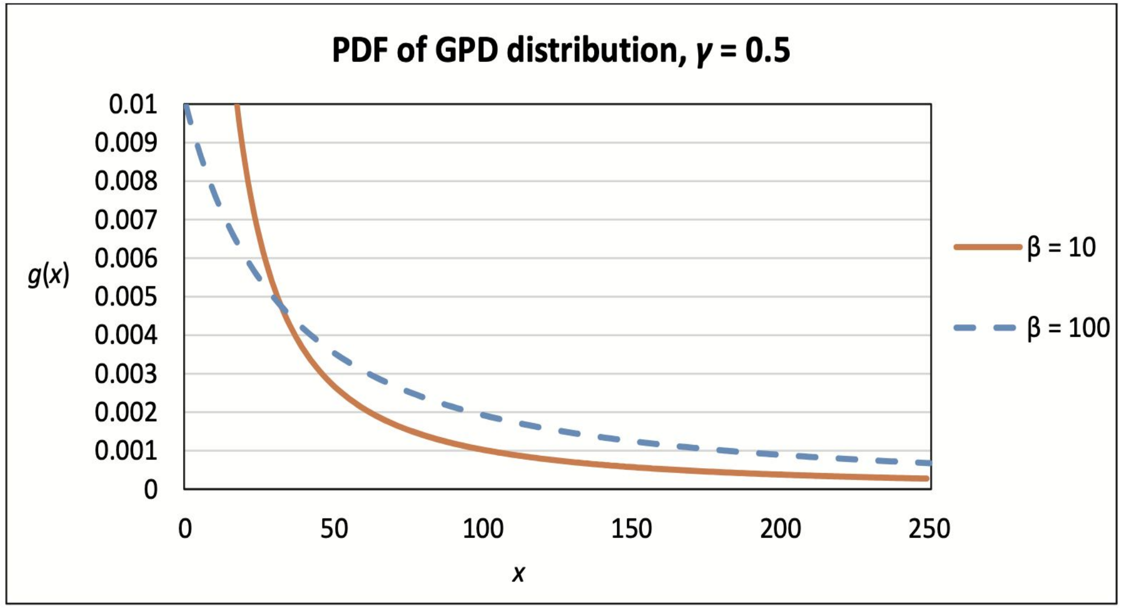
The generalised Pareto distribution is a two-parameter distribution with CDF:

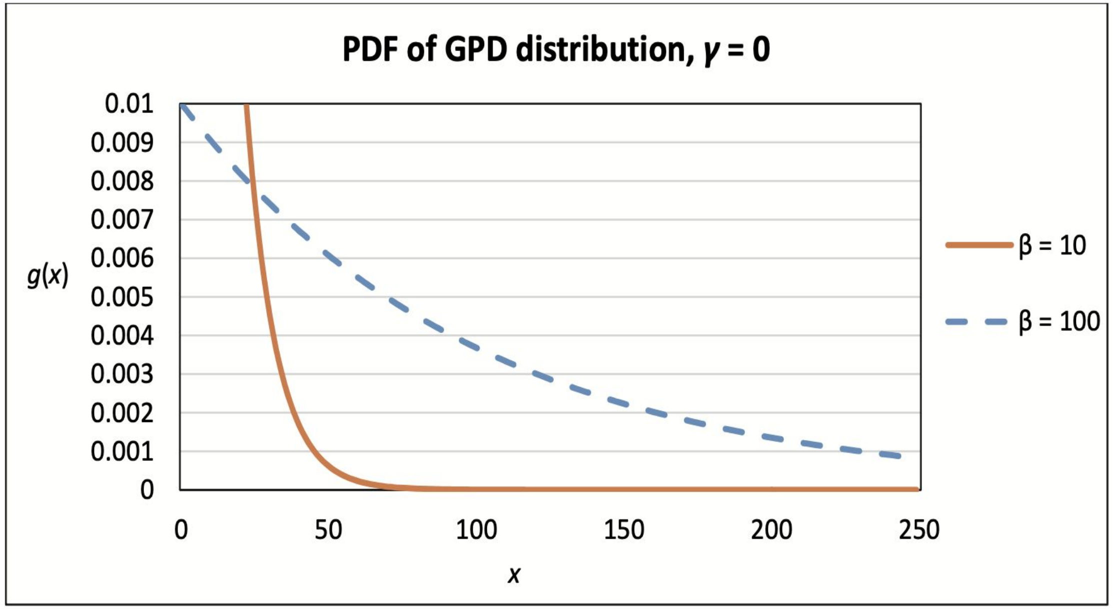
$$G(x)=\begin{cases}1-\left(1+\dfrac{x}{\gamma\beta}\right)^{-\gamma} & \gamma\neq 0\\ 1-\exp\left(-\dfrac{x}{\beta}\right) & \gamma=0\end{cases}$$

3. Generalised Pareto Distribution

PDF of the GPD distribution:

3. Generalised Pareto Distribution





3. Generalised Pareto Distribution

Key Advantage of the GPD

- The General Pareto Distribution has the advantage that is uses a larger part of the data and models all the large claims above the threshold, not just single higher value.

4. Measures of Tail Weights

➢ Existence of Moments (raw and central)

If more number of non central moments exist than that distribution has a lighter tail

Eg Pareto (α, λ) $\rightarrow moment\ exist$ if K < α

$E(X^K) = \sum X^k P(x)$

higher α $\rightarrow more\ moments$ $\rightarrow lighter\ tail$

lower α $\rightarrow less\ moments$ $\rightarrow fat\ tails$

Eg Ga(α, β)

4. Measures of Tail Weights

- Limiting Density Ratio

$$\lim_{x\to\infty} \frac{f_{X_1}(x)}{f_{X_2}(x)}$$

$= \infty \rightarrow$ numerator has heavier tail

$= 0 \rightarrow$ denominator has a heavier tail

Compare it for a Pareto distribution with α = 2 and α = 3 with the same λ

4. Measures of Tail Weights

Consider the $Gamma(0.5, 0.005)$ and $Pa(4, 300)$ distributions, both of which have a mean of 100 and a variance of 20,000.

Use the 'limiting density ratio' method to compare these two distributions.

4. Measures of Tail Weights

- Hazard Rate

- <u>Example – for mortality-</u>

f(t) = $tP_x * \mu_{x+t}$

F(t)= tq_x $\quad\quad$ $\mu_{x+t} = \frac{f(t)}{1-F(t)}$

$$h(x) = \frac{f(x)}{1-F(x)}$$

1. Hazard is an increasing function of x
   Tail will be lighter as less people reach higher ages
2. Hazard is a decreasing function of x
   Heavier Tail as more people reach higher ages

4. Measures of Tail Weights

(i) Determine the hazard rate for:

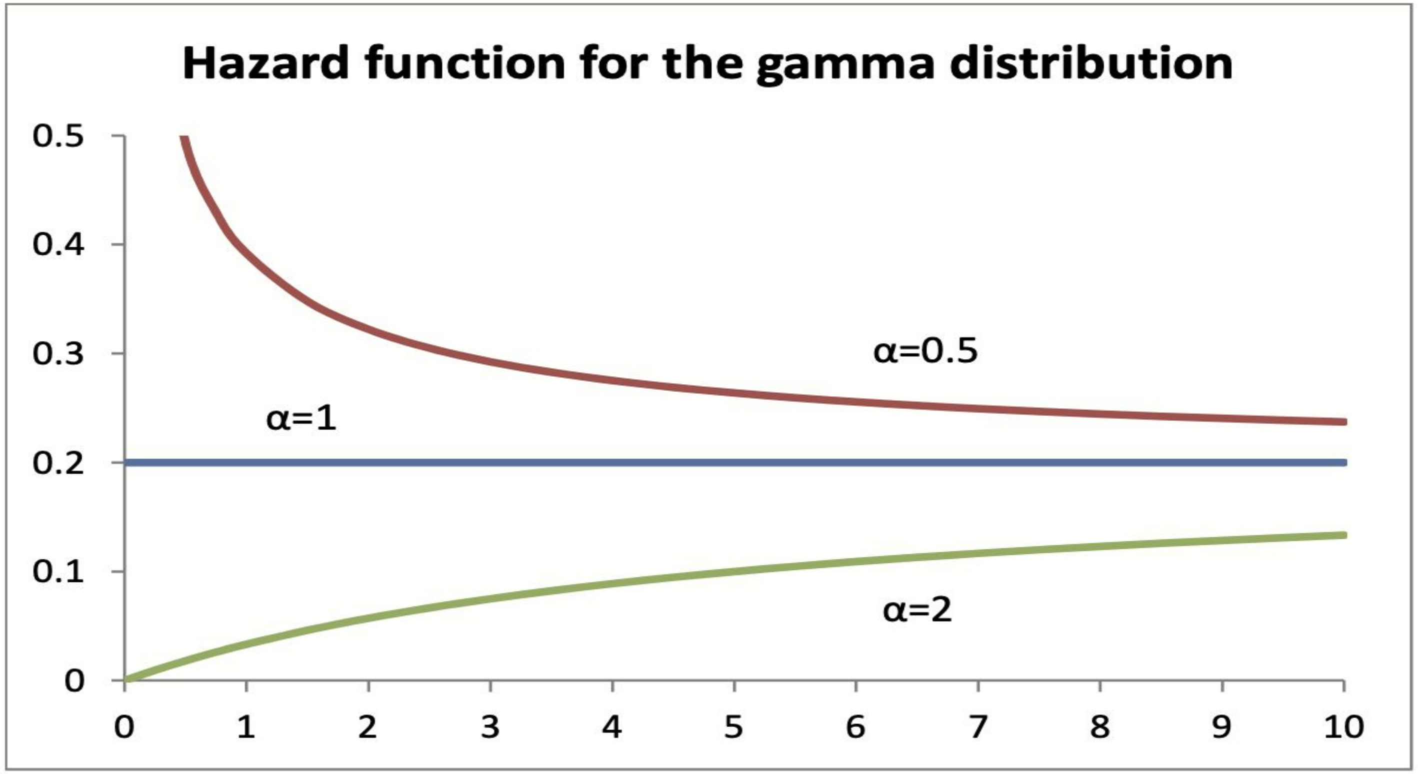
(a) the $Exp(\lambda)$ distribution

(b) the $Pa(\alpha,\lambda)$ distribution.

(ii) Comment on the differences between these hazard rates.

4. Measures of Tail Weights

For a Gamma distribution with same λ



4. Measures of Tail Weights

- Mean Residual Time

$$e(x) = \frac{\int_x^{\infty}(1-F_y(Y).dy}{1-F(x)}$$

If MRT is an increasing function of x ; fatter tail
If MRT is a decreasing function of x ; lighter tail

4. Measures of Tail Weights

(i) Determine the mean residual life for:

(a) the $Exp(\lambda)$ distribution

(b) the $Pa(\alpha, \lambda)$ distribution where $\alpha > 1$.

(ii) Comment on the behaviour of these functions.

4. Measures of Tail Weights

(i) Determine the mean residual life for:

(a) the $Exp(\lambda)$ distribution

(b) the $Pa(\alpha, \lambda)$ distribution where $\alpha > 1$.

(ii) Comment on the behaviour of these functions.

4. Measures of Tail Weights

(i) Show that:

$$\int_x^\infty e^{-3y^{\frac{1}{2}}}\,dy = \frac{2}{9}P\left(\chi_4^2 > 6x^{\frac{1}{2}}\right) \qquad [5]$$

Hint: use the substitution $u = 3y^{\frac{1}{2}}$ *and transform the integrand into the PDF of the* $Gamma(2,1)$ *distribution.*

(ii) Hence deduce an expression involving a chi-squared probability for the mean residual life for the $W\left(3, \frac{1}{2}\right)$ distribution. [2]

(iii) By calculating the values of mean residual life function when $x = 1$, $x = 4$ and $x = 9$, determine whether the mean residual life of the $W\left(3, \frac{1}{2}\right)$ distribution is an increasing or decreasing function of x. [2]

4. Measures of Tail Weights