Matβlas

Underwriting and Actuarial Consultancy, Training and Research

Demystifying Basic Statistical Concepts Used In The Insurance Industry

Prepared and presented by
Ana J. Mata, PhD, ACAS
Managing Director and Actuary
London, 2018

Disclaimer: The content of this note is theoretical in nature and exclusively prepared for training purposes. It should only be used as a reference for those attending or that have attended the presentation on *Demystifying Basic Statistical Concepts Used In The Insurance Industry*. No part of the content of this course material constitutes actuarial advice for any company or organization. All worked examples shown in this note were derived from hypothetical assumptions only and do not reflect any company's rates, rating factors, benchmark factors, loss experience or reporting processes. Any similarity between these examples and any insurance company is coincidental only.

Learning Objectives

The need for a statistical and probability framework in insurance and reinsurance.

Understanding basic statistical concepts widely used in insurance pricing, reserving and capital modelling.

Matβlas

Key Topics Covered

Probability models:
Single answer vs possible outcomes

The expected value of a random quantity

Quantifying risk

Applications to Insurance

Matβlas

Insurance Rating and Profitability

Statistical models are used to forecast quantities that are random (not fixed)

What are rating models in insurance forecasting in order to achieve a certain level of profit?

Profit = **Premium - Claims - Expenses**

What Is the Cost of an Insurance Policy?

The cost of a policy	Covered events	Probabilities
Cost of claims		
Unknown	Exposure	Risk

When a policy is sold: The *Expected* Claim Cost

Matβlas

The Expected Claim Cost

"Expected" means on average in the future

The average claim cost of a policy in the future year and includes

Number of events	Cost of each event
Frequency	**Severity**

Even if they have not been historically observed

Matβlas

Events – Experience - Forecast

Random events

- Uncertain frequency and severity
- True probabilities unknown

Claims experience

- Events happened
- Actual claims cost generated by events

Estimated claims

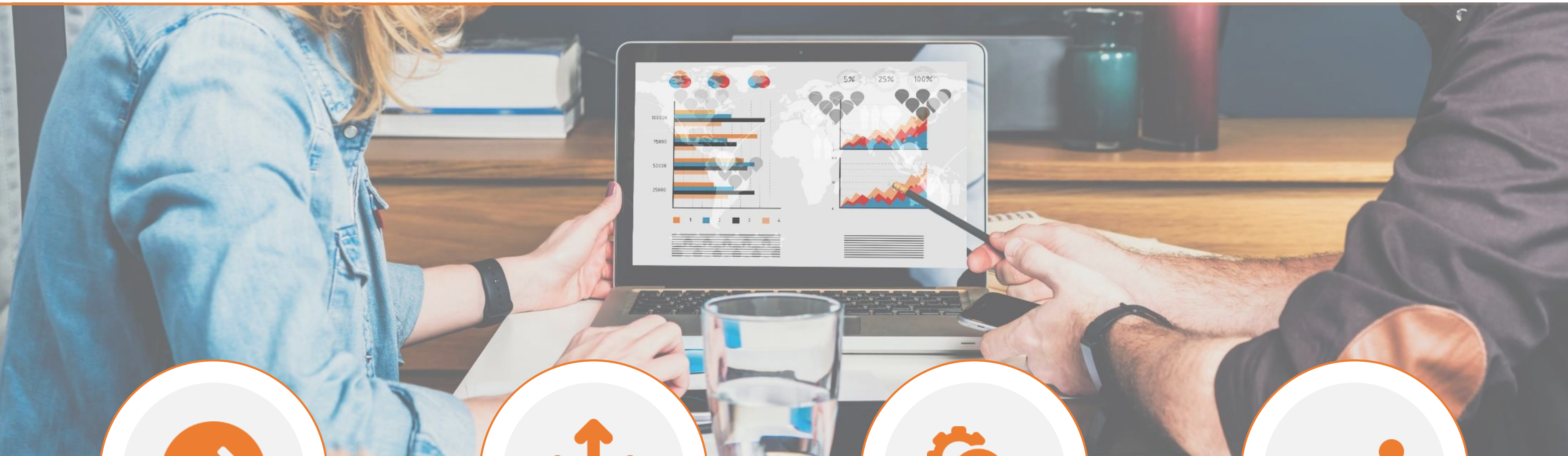
- Estimated frequency and severity
- Estimated premium
- Estimates have errors

Data are used to fit models, models are used to forecast

Matβlas

Probability Models

Probability Models



Possible outcomes are random

Possible (financial) outcomes are known (but not exact one)

Each possible outcome has a probability (adds to 100%)

Probability models allow us to calculate any quantity of interest

Basic Example of Probability Model

Illustrative example

Throwing a die

6 Possible Outcomes

1, 2, 3, 4, 5, 6

All outcomes have the same probability

1/6

What is the probability distribution of the random number?

Possible outcome	Probability of outcome	Probability ≤ outcome	Probability > outcome
1	1/6 = 16.66%	1/6 =16.67%	5/6 = 83.33 %
2	1/6 = 16.66%	2/6 = 33.33%	4/6 = 66.67%
3	1/6 = 16.66%	3/6 = 50.00%	3/6 = 50.00%
4	1/6 = 16.66%	4/6 = 66.67%	2/6 = 33.33%
5	1/6 = 16.66%	5/6 = 83.33%	1/6 = 16.67%
6	1/6 = 16.66%	6/6 = 100.00%	0%

Key Functions of Random Variables

Incremental: probability of taking an exact value x

$f(x)$ = Pr (Loss = x)

Cumulative probability: less than or equal to x

F(x) = Pr (Loss ≤ x)

Probability distribution

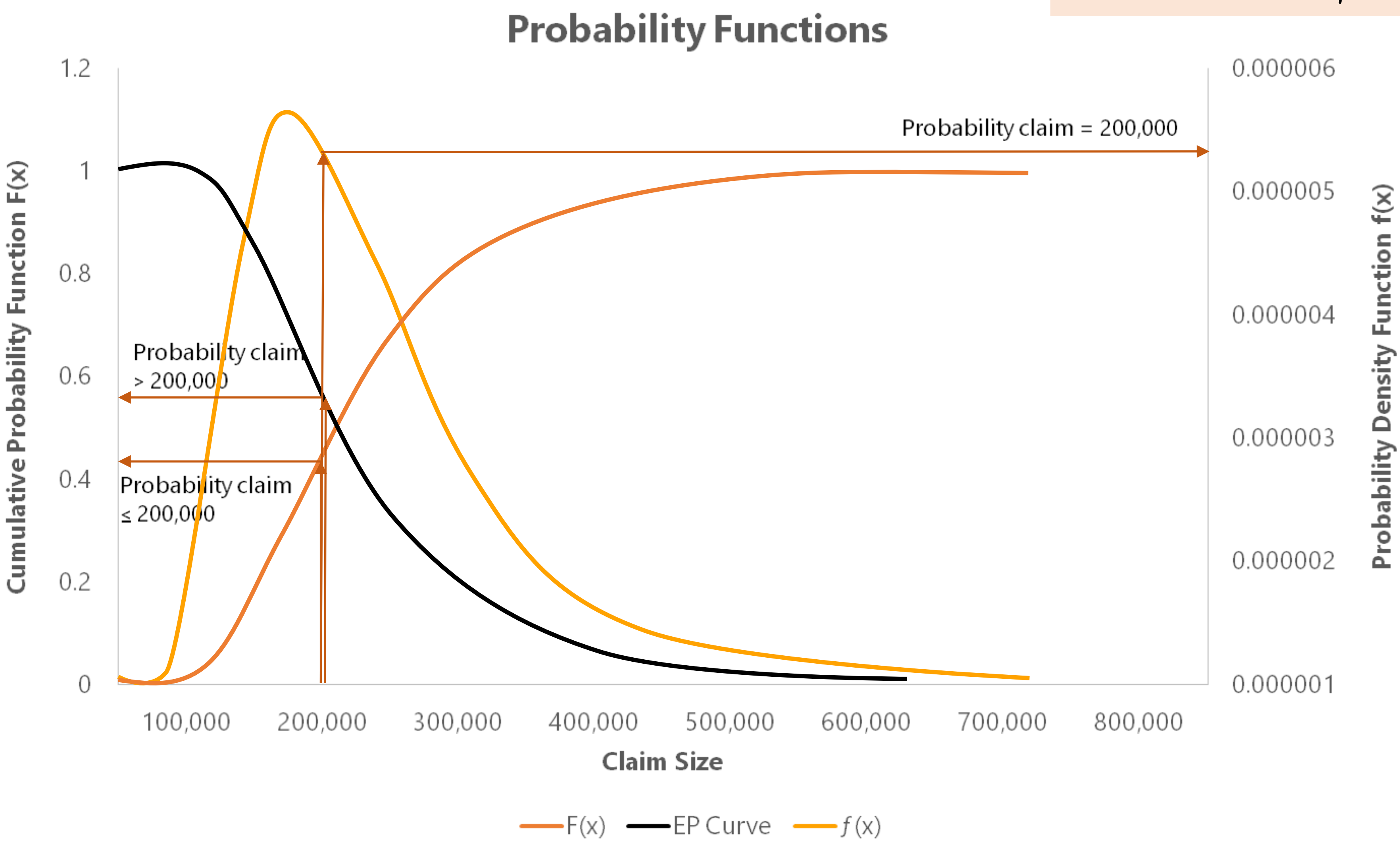
Cumulative probability: greater than x

$EP(x)$ = Pr (Loss > x) = 100% - Pr (Loss ≤ x)

Exceedance probability (use in catastrophe modelling)

Probability Functions of a Random Variable

Illustrative example



Matβlas

The Expected Value Of A Random Quantity

The Average vs. The Expected Value

Throwing a die

6 Possible Outcomes

1, 2, 3, 4, 5, 6

All outcomes have the same probability

1/6

Observations:

Throw the die 12 times and the following results are observed

1, 3, 4, 1, 2, 5, 3, 1, 3, 2, 5, 5

Observed average = 2.917

But we did not observed a 6

6 is a possible outcome

The average is underestimated

Matβlas

Average vs. Weighted Average

The average is simple the sum of all values divided by how many numbers are included in the sum.

$$Average = \frac{Sum\ all\ values}{Number\ of\ values}$$

The weighted average takes into account that each value has a different contribution (weight) to the average.

$$Wgt\ Average = \frac{Sum\ (value\ \times weight)}{Sum\ (weights)}$$

Matβlas

The Expected Value of a Random Quantity

Illustrative example

The expected value is the weighted average of all possible outcomes and the corresponding probabilities.

Expected value= ***Sum*** {possible outcome × its probability}

Across all possible loss outcomes

The expected value of the result of throwing a die

1 x 1/6 + 2 x 1/6 + 3 x 1/6 + 4 x 1/6 + 5 x 1/6 + 6 x 1/6 = 21/6 = 3.5

Note: the average or the expected value need not be one of the possible outcomes

Probabilities And The So Called 1-in-X Years

1-in-10 years, 1-in-20 years, …
very loosely used; very misunderstood.

The most common interpretation:

An event that has a probability of happening of 1-in-X years (the return period)

Frequency of an actual event happening in a period of time.

Frequency of claims or events within a portfolio.

The probability of cost of claims exceeding a certain amount: 1-in-10 chance that aggregate claims for the year will exceed £100m.

Matβlas

Frequency Models

Illustrative example

Frequency relates to the number of events that lead to losses/claims for a single policy or for a portfolio

Example: average frequency is 5%

"1 in 20 years"

<ins>On average</ins>, one such event occurs in a 20-year period; or

<ins>On average</ins>, on a portfolio of 20 policies, we <ins>expect</ins> one such claim each year

However, on any one year or policy there could be more than one such event per year (even though the probability of this is very low)

The following table shows a frequency probability distribution* with average of 5%:

No. Losses	Probability	No. losses x probability
0	95.12%	0
1	4.76%	0.0476
2	0.12%	0.0024
3	0.00%	0
4	0.00%	0
5	0.00%	0
	Expected frequency =	**0.05 (5%)**

***A Poisson distribution with average of 5% has been used**

Matβlas

Severity Models

Illustrative example

Once an event occurs, the severity is the loss generated by the event. The following table shows an example of a severity probability distribution:

Possible loss	Probability
250,000	42%
500,000	25%
1,000,000	14%
3,000,000	16%
5,000,000	3%

62% of the time the cost of the event will be £500k or less.
What is the average or expected cost per event with this probability model?

Matβlas

How Often vs. How Much?

Illustrative example

Common misconception: The **mean** or average is the mid point of the distribution: 50%-50% chance to each side.

Possible Loss	Probability	Loss x Probability
250,000	**42%**	**105,000**
500,000	25%	125,000
1,000,000	**14%**	**140,000**
3,000,000	16%	480,000
[illegible]	**3%**	**[illegible]**
Expected severity =		**[illegible]**

The **mode** is the most likely outcome

The **median** is the mid-point of the distribution (50% probability either side)

A skewed probability distribution is NOT symmetric around the mean

Matβlas

Skewness in General Insurance

FACT:

Insurance rates and premiums calculated based on average costs

FACT:

Insurance claims are skewed (Mean > Median)

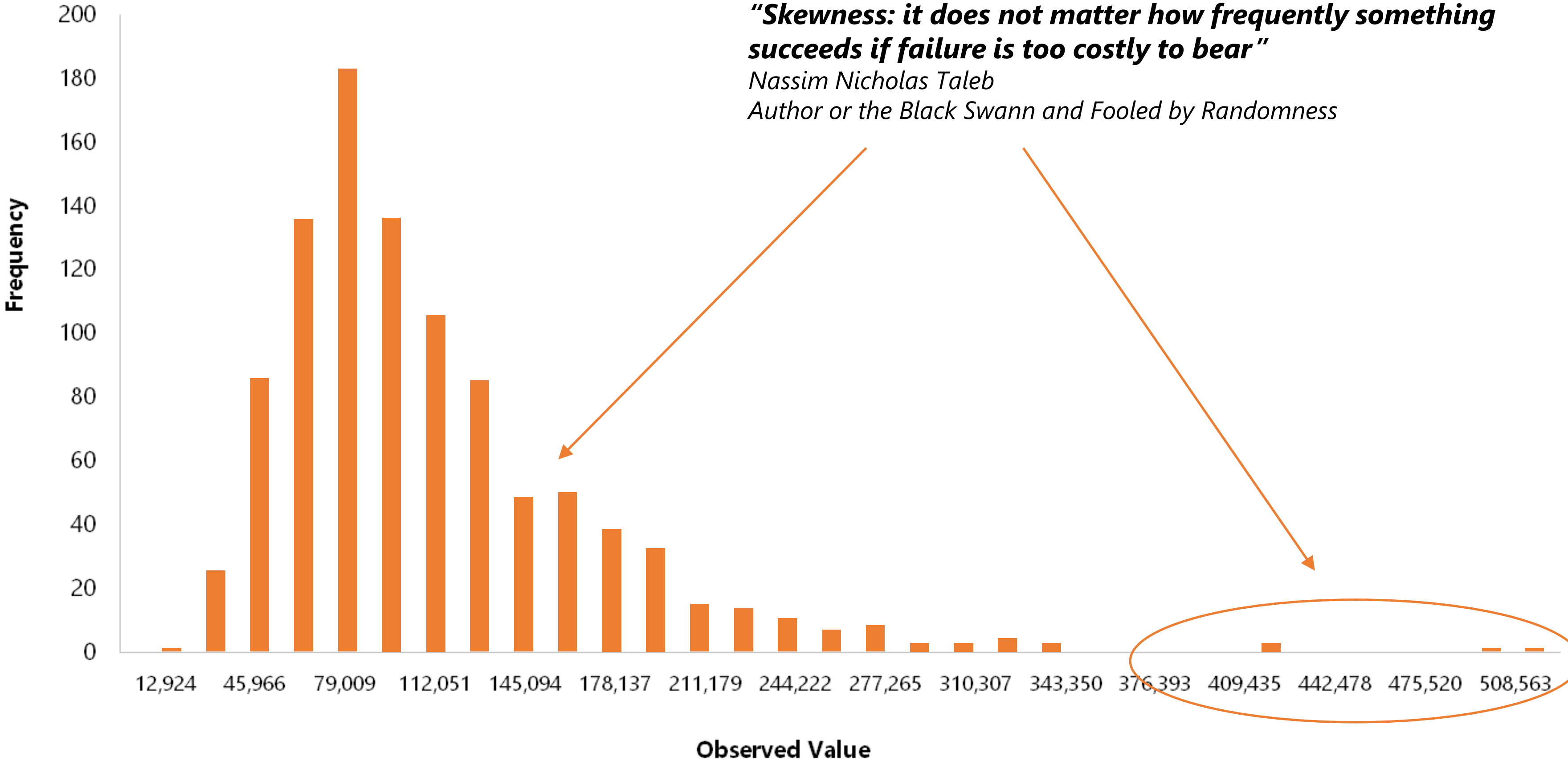
THEREFORE:

We have more than 50% chance of doing "better" than average (the mean)

ISN'T THIS GOOD NEWS FOR (RE)INSURERS?

Skewness in General Insurance

Illustrative example



"Skewness: it does not matter how frequently something succeeds if failure is too costly to bear"
Nassim Nicholas Taleb
Author or the Black Swann and Fooled by Randomness

Skewness is the length of the tail: how far are extreme possible outcomes

Matβlas

Skewness in General Insurance

Pricing

Expected cost of claims and loss ratio target.

Reserving

Expected cost of claims set aside as a reserve: actuarial best estimate of future payments.

Risk Capital

When claims are worse than expected; insufficient reserves (risk).

Matβlas

Quantifying Risk

Can You Accurately Price a Single Policy?

Illustrative example

1% (frequency)
Probability of a claim on any policy

£1m
Policy limit

Assume every event generates a full limit claim

Expected claim cost per policy = frequency x severity = 1% x £1M = £10,000

Premium charged = £20,000 Expected loss ratio = £10k/£20k = 50%

Claims experience will be 0 or £1m Loss ratio 0% or 5,000%

Is the price accurate?

Matβlas

Insurance Based on Volume and Years

Pricing is Based On

Illustrative example

Expected future average costs across policyholders and across years

Year	Premium	No. Losses	Annual Loss	Loss Ratio
1	20,000,000	11	11,000,000	55.00%
2	20,000,000	11	11,000,000	55.00%
3	20,000,000	20	20,000,000	100.00%
4	20,000,000	8	8,000,0000	40.00%
5	20,000,000	6	6,000,000	30.00%
6	20,000,000	13	13,000,000	65.00%
7	20,000,000	8	8,000,000	40.00%
8	20,000,000	5	5,000,000	25.00%
9	20,000,000	9	9,000,000	45.00%
10	20,000,000	13	13,000,000	65.00%

10-year average LR = 52%

Was the price accurate?

Matβlas

The Law of Large Numbers

1

True probabilities, outcomes and mean

- Cannot be accurately calculated
- Can be estimated (statistical models)

2

Actuarial methods are based on averages

3

Large volume of data more reliable to estimate averages

- % share of the market
- Number of years of experience

Matβlas

Key Functions of a Probability Distribution

Illustrative example

How are outcomes and probabilities used to measure risk?

Possible Claim	Probability of Exact Value f(x)	Probability Less Than F(x)	Probability of Greater Than EP(x) = 100% - F(x)
250,000	42%	42%	58%
500,000	25%	67%	33%
1,000,000	14%	81%	19%
3,000,000	16%	97%	3%
5,000,000	3%	100%	0%

Risk

Matβlas

Percentiles of a Probability Distribution

Illustrative example

Percentiles are NOT probabilities.

Percentiles are the possible outcomes associated with cumulative probabilities.

95th percentile: 95% of the time the random value will be less than or equal to the 95th percentile.

Possible Claim	Probability of Exact Value	Probability of less than or equal to	Probability of Greater Than
250,000	42%	42%	58%
500,000	25%	67%	33%
1,000,000	14%	81%	19%
3,000,000	**16%**	**97%**	**3%**
5,000,000	3%	100%	0%

3m is the 97th percentile

97% the of the time the claim will be 3m or less

The mean is the 81st percentile

Matβlas

The Value-at-Risk (VaR)

Illustrative example

Value-at-Risk is NOT a probability.

5% VaR is also called the 1-in-20 in the context of risk and capital

5% VaR: 5% of the time the random value will be greater than the 5% VaR (is a threshold)

Possible Claim	Probability of Exact Value	Probability of less than or equal to	Probability of Greater Than
250,000	42%	42%	58%
500,000	25%	67%	33%
1,000,000	14%	81%	19%
3,000,000	**16%**	**97%**	**3%**
5,000,000	3%	100%	0%

3m is the 3% VaR

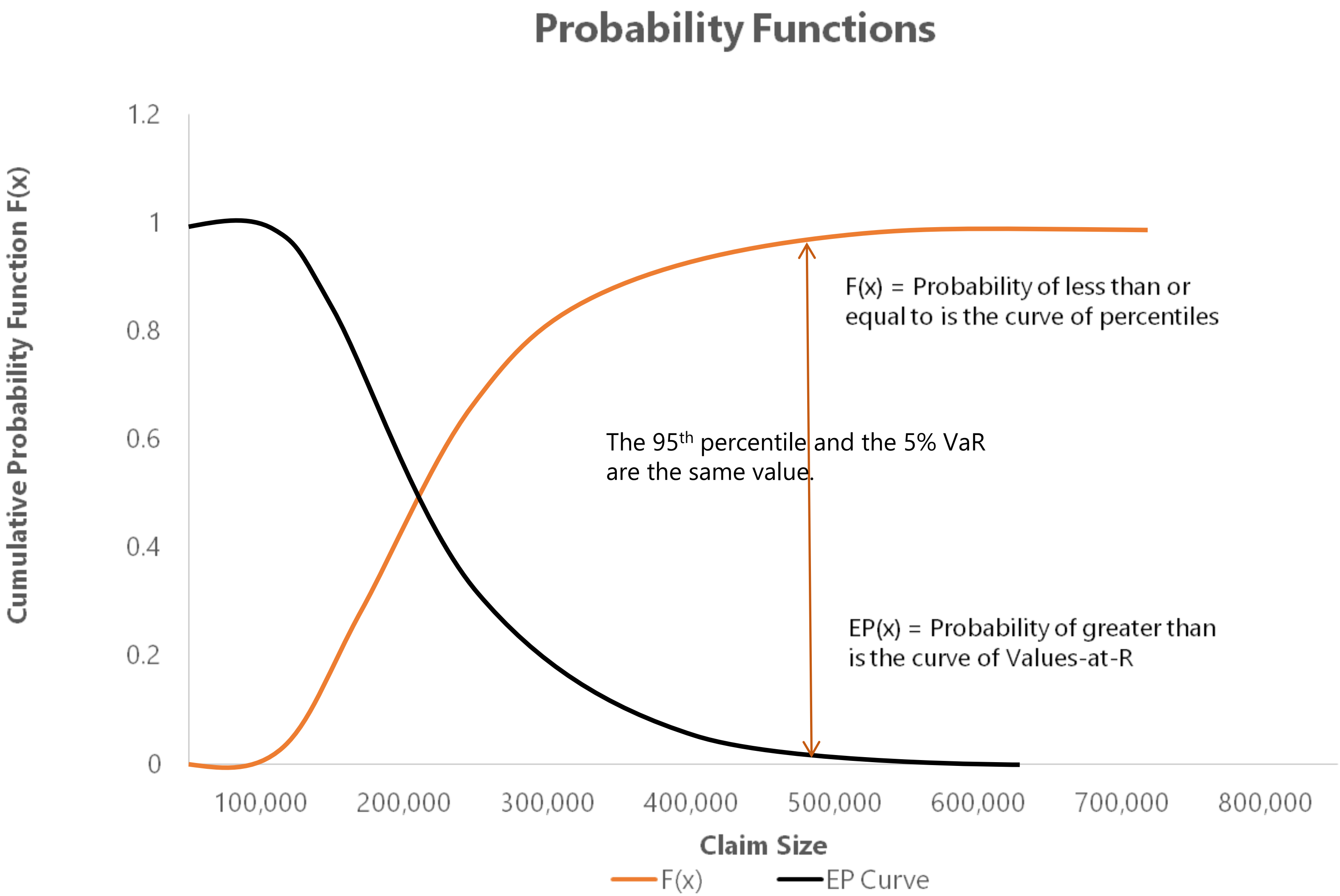
3% the of the time the claim will be greater than 3m

The mean is the 19% VaR

Matβlas

Percentiles and Value-at-Risk

Illustrative example



Matβlas

How Far From The Mean?

Illustrative example

1% (frequency)

Probability of a claim on any policy

£1m

Policy limit

Expected claim cost per policy = £10,000; Expected Loss Ratio = 50%

No. of claims	Prob. ≤ No. Claims (1 policy)
0	99%
1	100%
5	
10	
15	
20	

1 policy:

Reserves held: £10,000

Capital to cover 1 claim = £1m - £10k = £990,000

9,900% of reserves

Matβlas

How Far From The Mean?

Illustrative example

1% (frequency)

Probability of a claim on any policy

£1m

Policy limit

Expected claim cost per policy = £10,000; Expected Loss Ratio = 50%

No. of claims	Prob. ≤ No. Claims (1000 policy)
0	0%
1	0%
5	6.6%
10	58.3%
15	95.2%
20	99.9%

1,000 policies:

Reserves held: £10m

Capital to cover 20 claims = £20m - £10m = £10m

100% of reserves

Matβlas

The Standard Deviation

Used as a unit of “distance”

Always relative to the average

How many standard deviations is the 99th percentile (1-in-100) from the mean?

Confidence intervals around the mean:

+/- 2 standard deviations

Risk loads: a certain % of the standard deviation

Matβlas

A Relative Measure of Dispersion

The coefficient of variation (CoV) is widely used in insurance as a measure of volatility

$$\text{Coefficient of Variation (CoV)} = \frac{\textit{St. Deviation}}{\textit{Mean}}$$

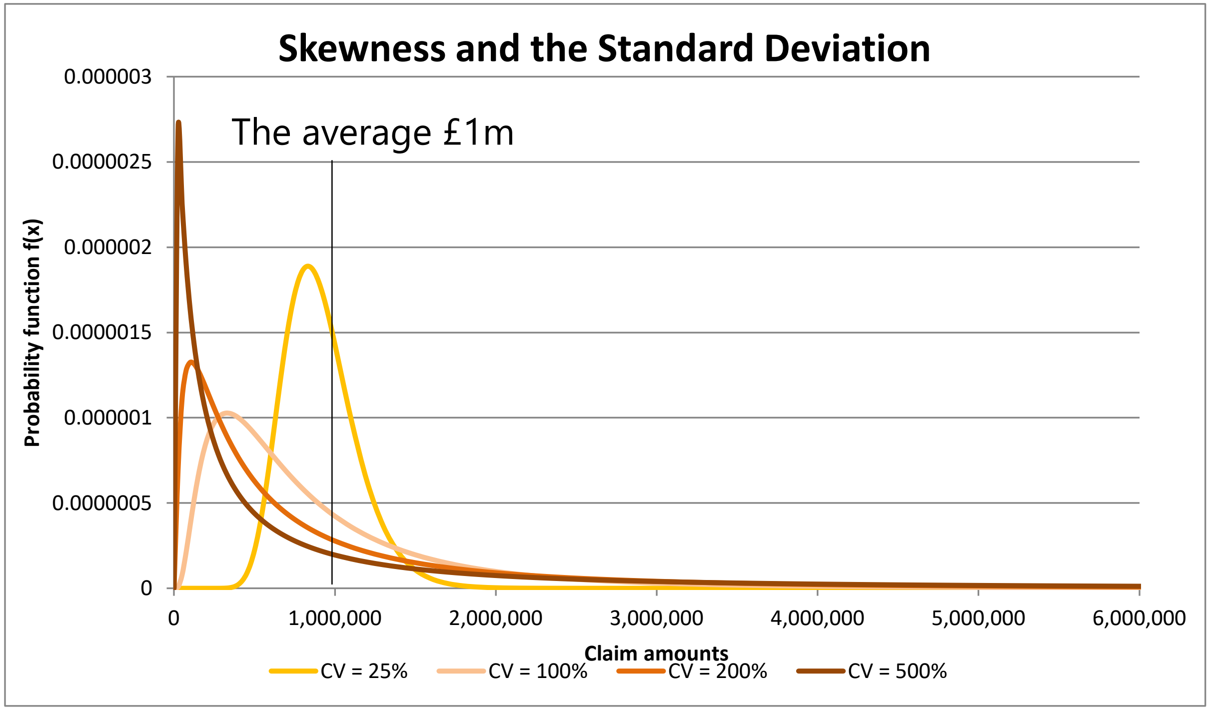
Measure of "dispersion" as a % of the mean

The smaller the CoV the smaller the dispersion and vice versa

Matβlas

Skewness And The Standard Deviation

Illustrative example



Average	£1,000,000			
CoV (%)	25%	100%	200%	500%
90th percentile (1-in-10)	1.21m	1.92m	2.70m	4.10m
95th percentile (1-in-20)	1.33m	2.60m	4.27m	7.90m
99th percentile (1-in-100)	1.57m	4.57m	10.14m	27.02m
99.5th percentile (1-in-200)	1.67m	5.63m	13.92m	42.39m

Matβlas

Capital Requirements Under Solvency II

- Include all risks: reserving, investment, credit, operational,…
- Time horizon: 12 months (not 200 years!).
- Risk tolerance: 1-in-200, hold risk capital to cover 99.5% of all possible outcomes. Leave only 0.5% chance of insolvency.
- Capital = Reserves + Risk Capital = 99.5th percentile.
- Simulate possible outcomes using statistical models (the mean and the standard deviation)
- Sort from largest (worst case) to smallest (best case). 1-in-200 threshold: 0.5% worse simulated outcomes exceed this value.

Capital Requirements Under SII

Illustrative example

- Portfolio of motor insurance
- £100m premium
- Avg. annual claims £68m (68% LR) and high volatility (>100%)
- 1,000 possible outcomes
- Worst 5 outcomes = 0.5% (1-in-200)

Capital required (99.5th)= £330.3m

Reserves held = £68m

Risk Capital = £330.3m - £68m = 262.3m

	Position highest to lowest	Annual claims (£m)	
Worst 10 scenarios	1000	£819.2	Worst 0.5%
	999	£615.3	
	998	£485.2	
	997	£381.2	
	996	£357.4	1-in-200
	995	£330.3	
	994	£322.8	
	993	£320.1	
	992	£317.0	
	991	£312.2	
	.	.	Average £68m
	.	.	
	.	.	
Best 10 scenarios	10	£8.0	
	9	£7.8	
	8	£7.6	
	7	£7.2	
	6	£6.9	
	5	£6.0	
	4	£5.7	
	3	£5.3	
	2	£3.9	
	1	£3.8	

Matβlas

Diversification – Pooling Risks

1

Diversification occurs when **independent risks** are pooled to reduce volatility around the mean

2

Diversification **does not** reduce the expected loss cost or average.

The average increases proportionally to the number of policies

3

Diversification **reduces** the potential financial downside when claims are worse than expected

Reduces the risk capital relative to premium volume

Risk (A + B) $\leq$ Risk (A) + Risk (B)

99.5th percentile of (A + B) $\leq$ 99.5th percentile of (A) + 99.5th percentile of (B)

Diversification and Risk Capital

Illustrative example

Property

Premium = 100m,

ULR = 65%, CoV = 150%

Casualty

Premium = 50m,

ULR = 75%, CoV = 70%

Marine

Premium = 35m,

ULR = 70%, CoV = 60%

Class	Premium	Claims Reserves	99.5th percentile (1-in-200)	Risk Capital	Risk Capital % Premium
Property	100m	65.0m	460.4m	395.4m	395%
Casualty	50m	37.5m	183.6m	146.1m	292%
Marine	35m	24.5m	112.8m	88.3m	252%

Combining 3 classes reduces volatility (lower 99.5th percentile)

Diversification benefit = reduction in capital 215m (34%)

Matβlas

Summarising

1

The Mean

- Single point
- Need not be a possible outcome
- Used for pricing and reserving
- Across policies/ across years

2

The Standard Deviation

- Volatility around the mean
- All possible outcomes
- How far from the mean (percentiles)

3

Percentiles and VaR

- Not related to frequency of events
- Risk tolerance: cumulative probabilities of annual claims

Wrap Up: Facts to Keep in Mind

Insurance is all about statistics

- Single answer vs. possible outcomes
- The average may never happen
- The fact that something has not happened does not mean it will not happen
- Things can happen more or less frequently than observed

Need to Understand Models and Their Limitations

- Facts vs. assumptions vs. opinions

Underwriting Management

- Diversification is key
- Balance volume vs. profitability

Matβlas