



Optimisation of Mitigations in a Multivariate Cyber Risk Model

M Bentley, A Stephenson, P Toscas and Z Zhu



Working on updates

Part 2 of 3: Installing features and drivers

50% complete

Don't turn off your computer, this will take a while

Doops, your important files are encrypted.

If you see this text, then your files are no longer accessible, because they have been encrypted. Perhaps you are busy looking for a way to recover your files, but don't waste your time. Nobody can recover your files without our decryption service.

We guarantee that you can recover all your files safely and easily. All you need to do is submit the payment and purchase the decryption key.

Please follow the instructions:

1. Send \$300 worth of Bitcoin to following address:

1Mz7153HMuxXTuR2R1t78mGSdzaAtNbBWx

2. Send your Bitcoin wallet ID and personal installation key to e-mail wowsmith123456@posteo.net. Your personal installation key:

74f296-2Nx1Gm-yHQRWr-S8gaN6-8Bs1td-U2DKui-ZZpKJE-kE6sSN-o8tizU-gUeUMa

If you already purchased your key, please enter it below.

Key: _



- “Tell me the cyber risk!”



- “Tell me the cyber risk!”
- “What do we spend the security budget on?!”



Current Approaches



Current Approaches



- Qualitative: e.g. colour coding, categorisation, flow charts.

Current Approaches



- Qualitative: e.g. colour coding, categorisation, flow charts.
- Quantitative: e.g. Gordon-Loeb model.

Gordon-Loeb Model



The economics of information security investment

[LA Gordon](#), [MP Loeb](#) - ACM Transactions on Information and System ..., 2002 - dl.acm.org

This article presents an economic model that determines the optimal amount to invest to protect a given set of information. The model takes into account the vulnerability of the information to a security breach and the potential loss should such a breach occur. It is ...

☆  Cited by 1200 [Related articles](#) [All 19 versions](#)

The economics of information security investment

[LA Gordon](#), [MP Loeb](#) - ACM Transactions on Information and System ..., 2002 - dl.acm.org

This article presents an economic model that determines the optimal amount to invest to protect a given set of information. The model takes into account the vulnerability of the information to a security breach and the potential loss should such a breach occur. It is ...

☆  Cited by 1200 [Related articles](#) [All 19 versions](#)



[Main page](#)
[Contents](#)
[Featured content](#)
[Current events](#)
[Random article](#)
[Donate to Wikipedia](#)

Article [Talk](#) [Reactions](#)

Gordon–Loeb model

From Wikipedia, the free encyclopedia

The **Gordon–Loeb** /ˈɡɔr-dən ˈlɒb/ **Model** is a mathematical investment level in information security.

Investing to protect company data involves a cost to generate profit. It does, however, serve to prevent a expensive it is to protect a specific set of data, with damaged or corrupted. To draft this model, the com

The economics of information security investment

[LA Gordon](#), [MP Loeb](#) - ACM Transactions on Information and System ..., 2002 - dl.acm.org

This article presents an economic model that determines the optimal amount to invest to protect a given set of information. The model takes into account the vulnerability of the information to a security breach and the potential loss should such a breach occur. It is ...

☆  Cited by 1200 [Related articles](#) [All 19 versions](#)



Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Article [Talk](#)

Gordon-Loeb model

From Wikipedia, the free encyclopedia

The **Gordon-Loeb** /ˈɡɔːr-dən ˈləʊb/ **Model** is a mathematical investment level in information security.

Investing to protect company data involves a cost to generate profit. It does, however, serve to prevent a expensive it is to protect a specific set of data, with damaged or corrupted. To draft this model the com



Search



Gordon-Loeb Model for Cybersecurity Investments

Gordon-Loeb Model



Gordon-Loeb Model



$$ENBIS(z) = [\nu - S(z, \nu)]\lambda t - z$$

where

Gordon-Loeb Model



$$ENBIS(z) = [\nu - S(z, \nu)]\lambda t - z$$

where

- z - spending on security,

Gordon-Loeb Model



$$ENBIS(z) = [\nu - S(z, \nu)]\lambda t - z$$

where

- z - spending on security,
- ν - base probability of successful attack, with no spending,

Gordon-Loeb Model



$$ENBIS(z) = [\nu - S(z, \nu)]\lambda t - z$$

where

- z - spending on security,
- ν - base probability of successful attack, with no spending,
- $S(z, \nu)$ - probability of successful attack, with spending z , for example $S(z, \nu) = \nu^{\alpha z + 1}$,

Gordon-Loeb Model



$$ENBIS(z) = [\nu - S(z, \nu)]\lambda t - z$$

where

- z - spending on security,
- ν - base probability of successful attack, with no spending,
- $S(z, \nu)$ - probability of successful attack, with spending z , for example $S(z, \nu) = \nu^{\alpha z + 1}$,
- λ - fixed loss given successful attack,

Gordon-Loeb Model



$$ENBIS(z) = [\nu - S(z, \nu)]\lambda t - z$$

where

- z - spending on security,
- ν - base probability of successful attack, with no spending,
- $S(z, \nu)$ - probability of successful attack, with spending z , for example $S(z, \nu) = \nu^{\alpha z + 1}$,
- λ - fixed loss given successful attack,
- t - probability of an attack occurring.

Annualised Loss Expectancy

= Annualised Rate of Occurrence $\times$ Single Loss Expectancy

Annualised Loss Expectancy

= Annualised Rate of Occurrence $\times$ Single Loss Expectancy

or

$$ALE = ARO \times SLE$$

Annualised Loss Expectancy

= Annualised Rate of Occurrence $\times$ Single Loss Expectancy

or

$$ALE = ARO \times SLE$$

From the 1970s (at least); boils down to *frequency* and *severity*.

Operational Risk (Op Risk)



Operational Risk (Op Risk)



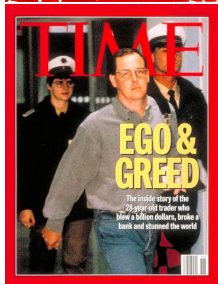
Operational Risk (Op Risk)



Operational Risk (Op Risk)



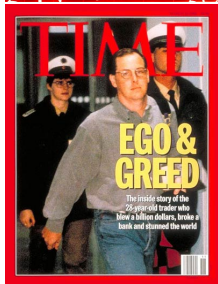
Operational Risk (Op Risk)



Operational Risk (Op Risk)



Operational Risk (Op Risk)



Op Risk Mathematical Model



Op Risk Mathematical Model



Two elements to an Op Risk model:

Op Risk Mathematical Model



Two elements to an Op Risk model:

- *Frequency* - How many incidents per year?

Op Risk Mathematical Model



Two elements to an Op Risk model:

- *Frequency* - How many incidents per year?
- *Severity* - How much damage occurs per incident?

Familiar Technology



Familiar Technology



Just fancy terminology for a garden variety compound Poisson process:

Familiar Technology



Just fancy terminology for a garden variety compound Poisson process:

$$L(t) =$$

Familiar Technology



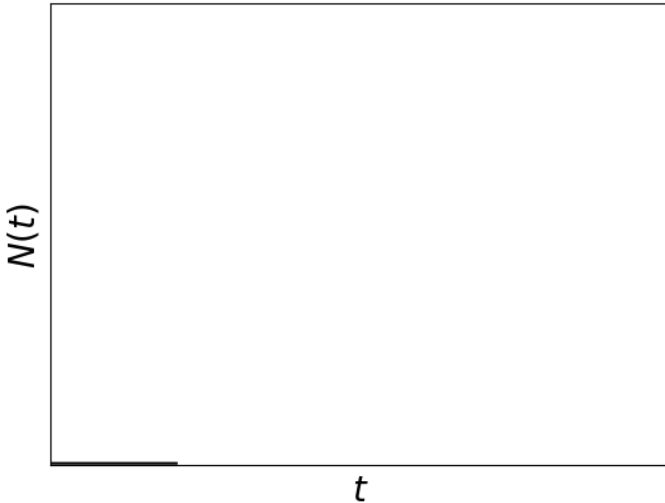
Just fancy terminology for a garden variety compound Poisson process:

$$L(t) = \sum_{i=1}^{N(t)}$$

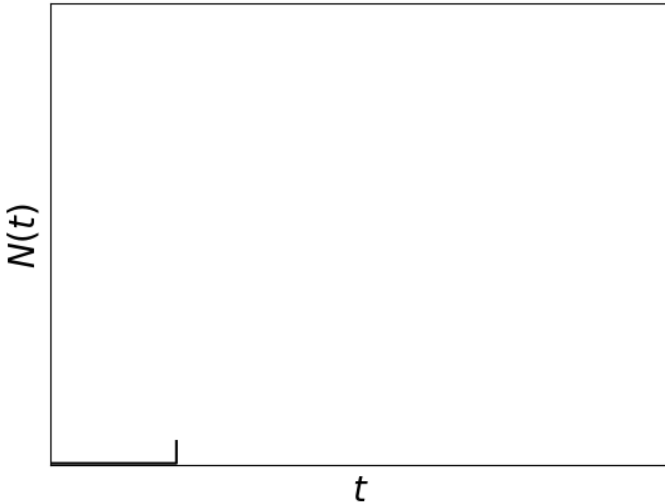
Just fancy terminology for a garden variety compound Poisson process:

$$L(t) = \sum_{i=1}^{N(t)} X_i$$

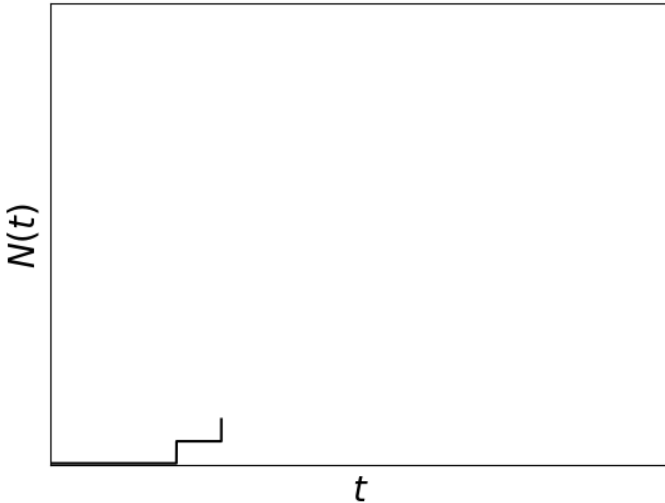
Poisson Process



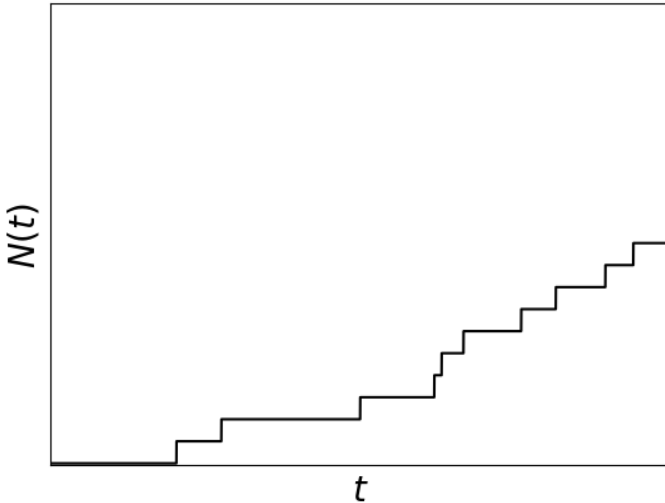
Poisson Process



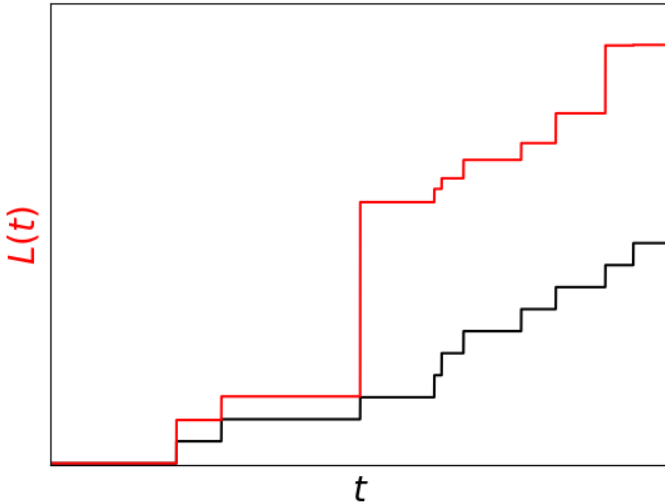
Poisson Process



Poisson Process



Compound Poisson Process



Mitigations by Thinning



Mitigations by Thinning



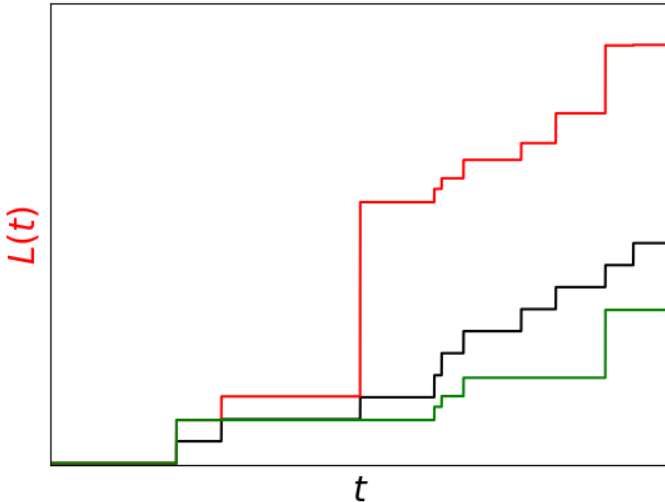
- If each arrival “succeeds” independently of other arrivals with probability q , the loss process is still compound poisson with intensity $q\lambda$.

Mitigations by Thinning



- If each arrival “succeeds” independently of other arrivals with probability q , the loss process is still compound poisson with intensity $q\lambda$.
- Idea: As per Gordon-Loeb, make q a function of spending on security, z .

Thinned Compound Poisson Process





- “It needs to handle at least four types of damage!”



Multivariate is easy!



- Easy! Set $j = 1, \dots, J$ as different types of damage / attack, then

Multivariate is easy!



- Easy! Set $j = 1, \dots, J$ as different types of damage / attack, then

$$L_j(t) = \sum_{i=1}^{N_j(t)} X_i^j$$

Multivariate is easy!



- Easy! Set $j = 1, \dots, J$ as different types of damage / attack, then

$$L_j(t) = \sum_{i=1}^{N_j(t)} X_i^j$$
$$L(t) = \sum_{j=1}^J L_j(t)$$

Multivariate is easy!



- Easy! Set $j = 1, \dots, J$ as different types of damage / attack, then

$$L_j(t) = \sum_{i=1}^{N_j(t)} X_i^j$$
$$L(t) = \sum_{j=1}^J L_j(t)$$

- Sum of compound Poisson is compound Poisson.

Multivariate is easy!

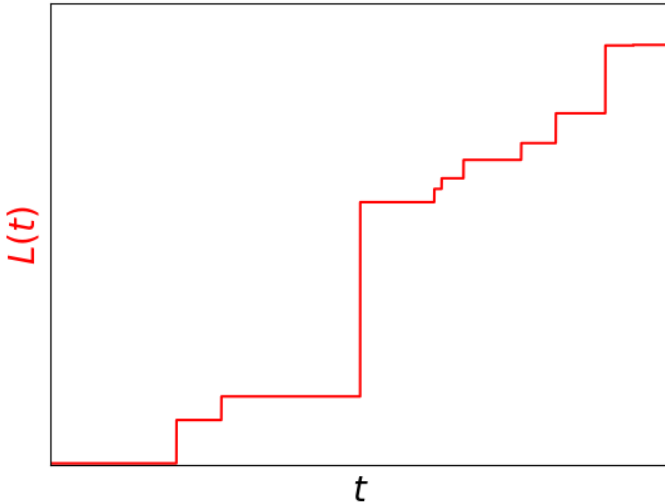


- Easy! Set $j = 1, \dots, J$ as different types of damage / attack, then

$$L_j(t) = \sum_{i=1}^{N_j(t)} X_i^j$$
$$L(t) = \sum_{j=1}^J L_j(t)$$

- Sum of compound Poisson is compound Poisson.
- Efficient numerical techniques to find distribution of $L(t)$, in particular using FFT.

Lots of Independence



Use a Copula



Use a Copula



- As many different ways to add dependence as there are forms of independence.

Use a Copula



- As many different ways to add dependence as there are forms of independence.
- Simple method: use a (Gaussian) copula function.

Use a Copula



- As many different ways to add dependence as there are forms of independence.
- Simple method: use a (Gaussian) copula function.

$$L(t) = \sum_{j=1}^J \sum_{i=1}^{N_j(t)} X_i^j$$

Not without controversy!



Not without controversy!



FELIX SALMON BUSINESS 02.23.09 12:00 PM

RECIPE FOR DISASTER: THE FORMULA THAT KILLED WALL STREET

$$\Pr[T_A < 1, T_B < 1] = \Phi_2(\Phi^{-1}(F_A(1)), \Phi^{-1}(F_B(1)), \gamma)$$

Here's what killed your 401(k) *David X. Li's Gaussian copula function as first published in 2000.*

Not without controversy!



FELIX SALMON BUSINESS 02.23.09 12:00 PM

RECIPE FOR DISASTER: THE FORMULA THAT KILLED WALL STREET



Not without controversy!



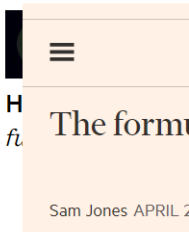
FELIX SALMON BUSINESS

RECIPE THAT KI

Extremes (2006) 9: 3–20
DOI 10.1007/s10687-006-0015-x

Copulas: Tales and facts

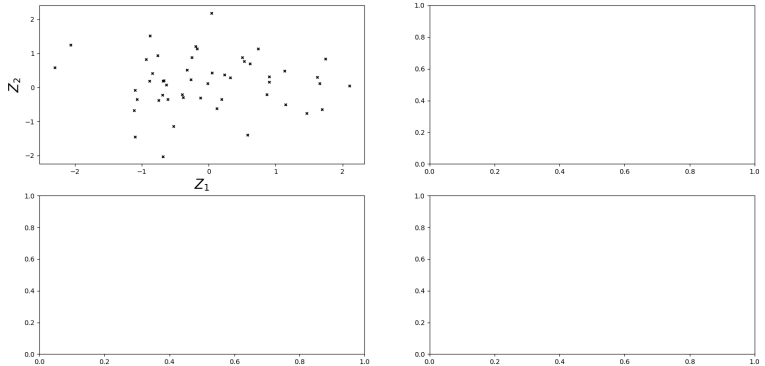
Thomas Mikosch



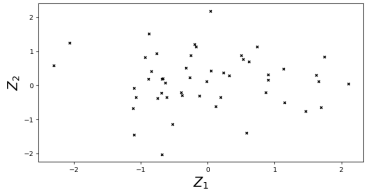
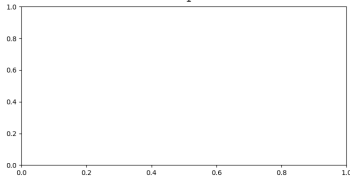
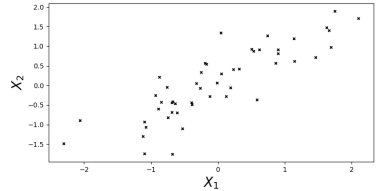
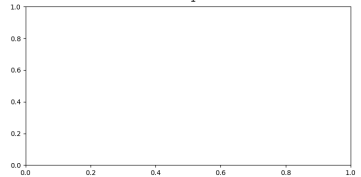
My main concern is that this very simple concept might be something like the emperor's new clothes because it promises to solve all problems of stochastic dependence but it falls short in achieving the goal.

one needs less than 10 minutes to understand the fundamentals of copulas, but many years of studies in order to get an idea of a genuine stochastic process.

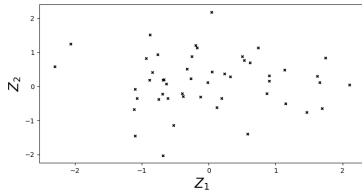
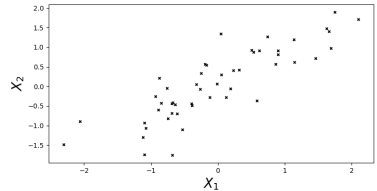
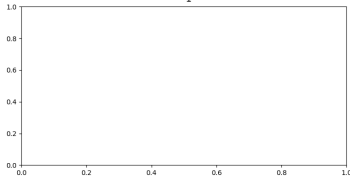
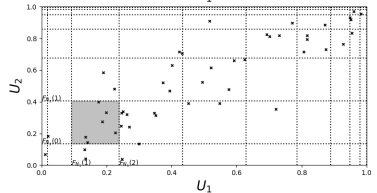
Simulation with a Gaussian Copula



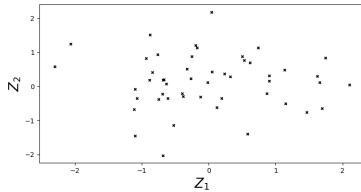
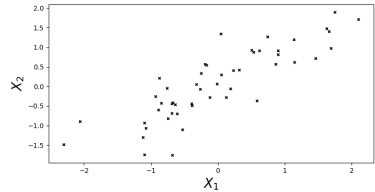
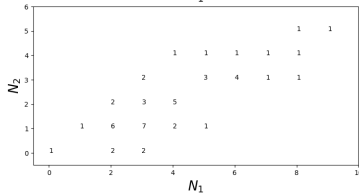
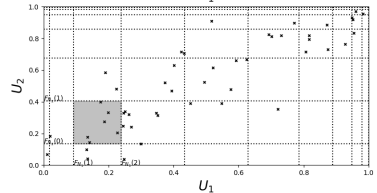
Simulation with a Gaussian Copula



Simulation with a Gaussian Copula



Simulation with a Gaussian Copula



Closed form Mean and Variance



Closed form Mean and Variance



$$\mathbb{E}[L] = \sum_{j=1}^J E[N_j] E[X_1^j]$$

Closed form Mean and Variance

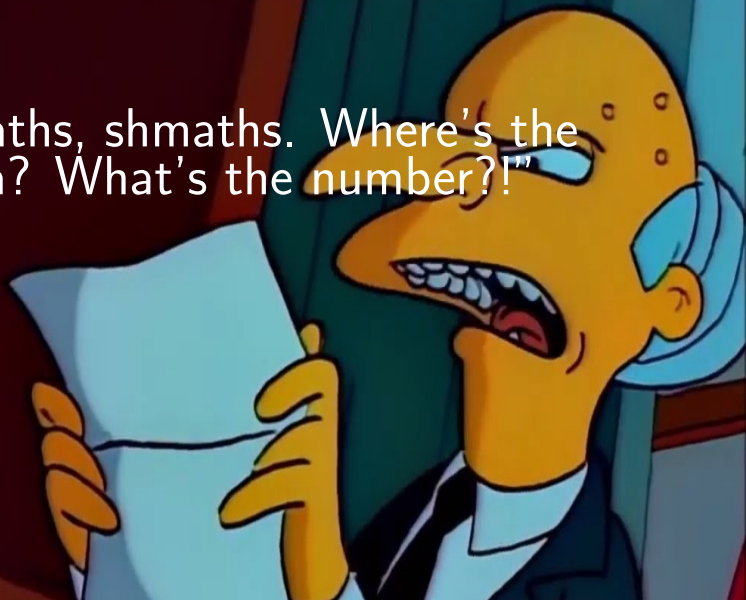


$$\mathbb{E}[L] = \sum_{j=1}^J \mathbb{E}[N_j] \mathbb{E}[X_1^j]$$

$$\begin{aligned} \text{Var}(L) = & \sum_{j=1}^J \left\{ \mathbb{E}[N_j] \text{Var}(X_1^j) + \text{Var}(N_j) \left(\mathbb{E}[X_1^j] \right)^2 \right\} \\ & + \sum_{i=1}^J \sum_{j=1}^J \rho_{ij} \mathbb{E}[X_1^i] \mathbb{E}[X_1^j] \sqrt{\text{Var}(N_i) \text{Var}(N_j)} 1(i \neq j) \end{aligned}$$



- “Maths, shmaths. Where’s the data? What’s the number?!”



(Free, specific) Data is scarce



(Free, specific) Data is scarce



- Data is out there! But nothing that is free and suitable.

(Free, specific) Data is scarce



- Data is out there! But nothing that is free and suitable.
- Free, unsuitable: Privacy Rights Clearinghouse, (DataLossDB), Veris Community Database, OWASP/WASC Web Hacking Incident Database, Hackmageddon.

(Free, specific) Data is scarce

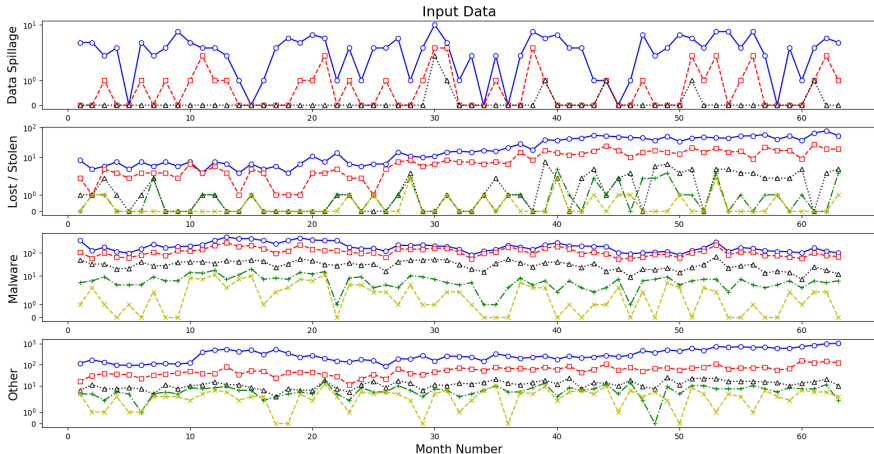


- Data is out there! But nothing that is free and suitable.
- Free, unsuitable: Privacy Rights Clearinghouse, (DataLossDB), Veris Community Database, OWASP/WASC Web Hacking Incident Database, Hackmageddon.
- Not free, suitable: SAS OpRisk Global Data, Cyber Risk Analytics, Cyber Loss Data by Advisen.

Some data in Kuypers et al 2016, Pate-Cornell et al 2018



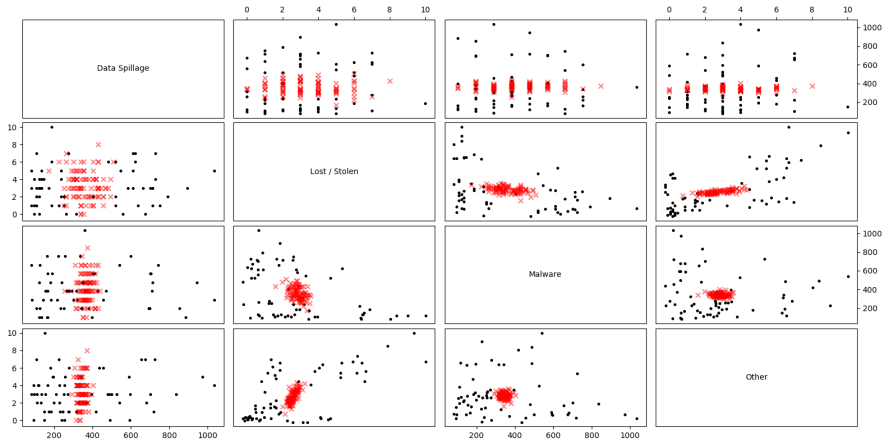
Some data in Kuypers et al 2016, Pate-Cornell et al 2018



Model should work...



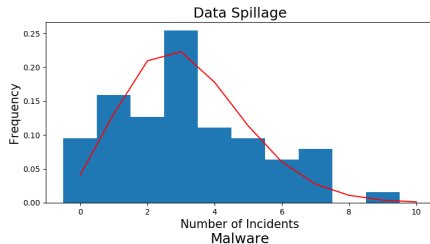
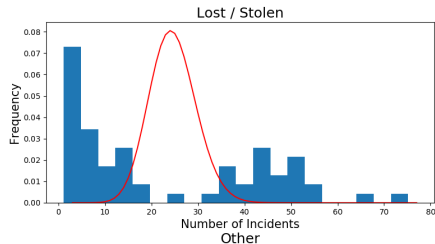
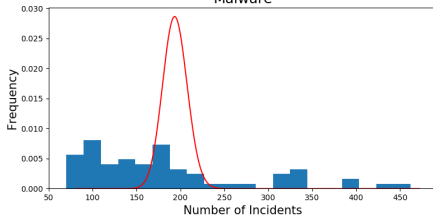
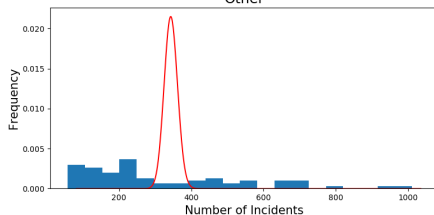
Model should work...



Marginals are wrong!



Marginals are wrong!



Use Negative Binomials



Use Negative Binomials



- Poisson has $E[N(t)] = \text{Var}(N(t))$ but the data is *overdispersed*.

Use Negative Binomials



- Poisson has $E[N(t)] = \text{Var}(N(t))$ but the data is *overdispersed*.
- Use Negative Binomial - basically Poisson, but intensity now has a Gamma $\Gamma(\alpha, \beta)$ distribution.

Use Negative Binomials

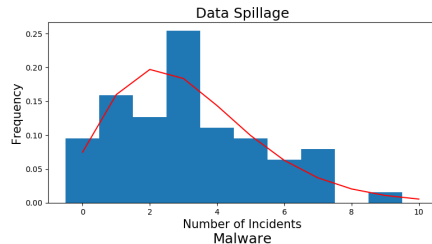
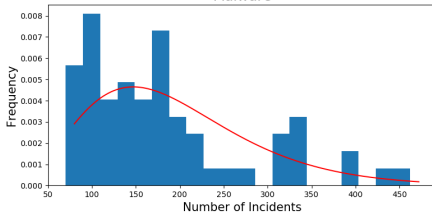
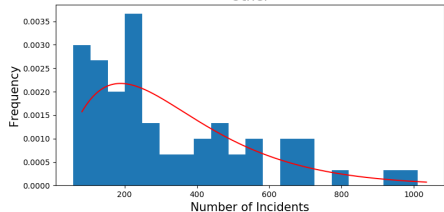


- Poisson has $E[N(t)] = \text{Var}(N(t))$ but the data is *overdispersed*.
- Use Negative Binomial - basically Poisson, but intensity now has a Gamma $\Gamma(\alpha, \beta)$ distribution.
- There's extra bookkeeping to deal with the thinning probabilities q_j , but it all works out.

Marginals are Better



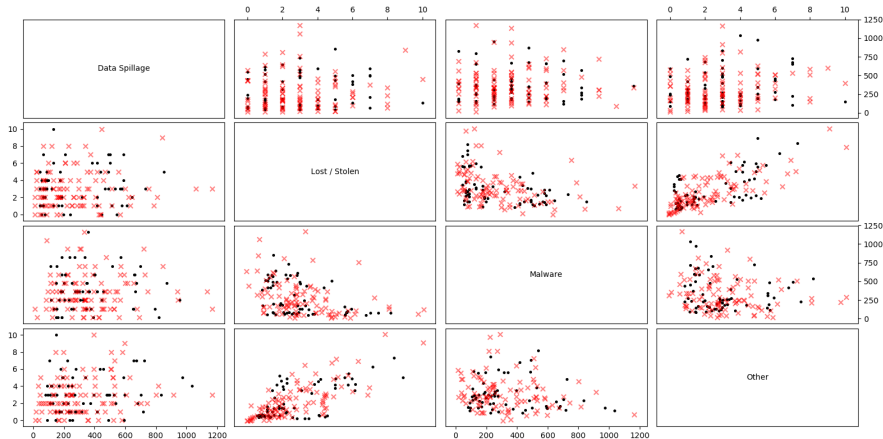
Marginals are Better



Multivariate Counts are Plausible



Multivariate Counts are Plausible



What about severity?



What about severity?



- Make assumptions on cost per hour of each incident.

What about severity?



- Make assumptions on cost per hour of each incident.
(Reasonable, but I'm being sneaky)

What about severity?



- Make assumptions on cost per hour of each incident.
(Reasonable, but I'm being sneaky)
- A very rough method to get distribution of severities.

What about severity?

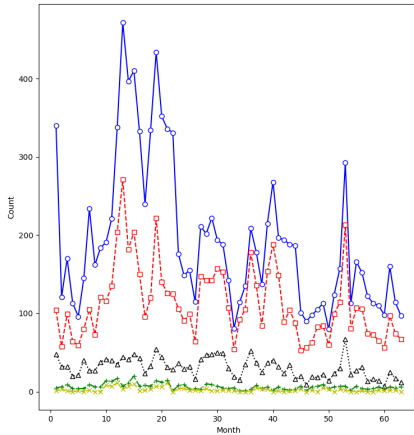
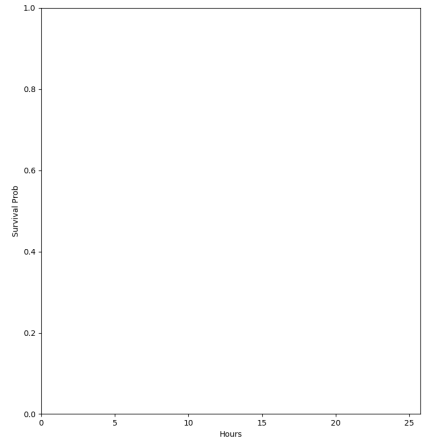


- Make assumptions on cost per hour of each incident. (Reasonable, but I'm being sneaky)
- A very rough method to get distribution of severities. (Quite primitive!)

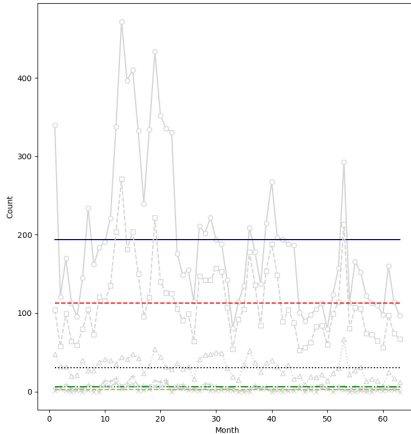
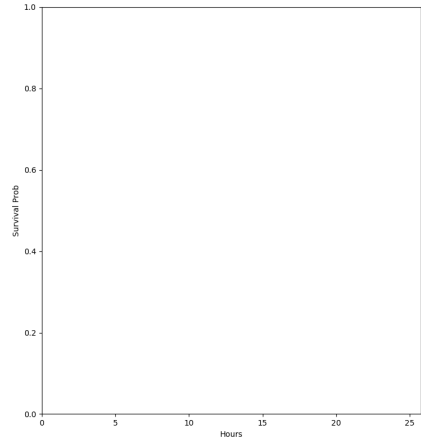
Rough method for severities



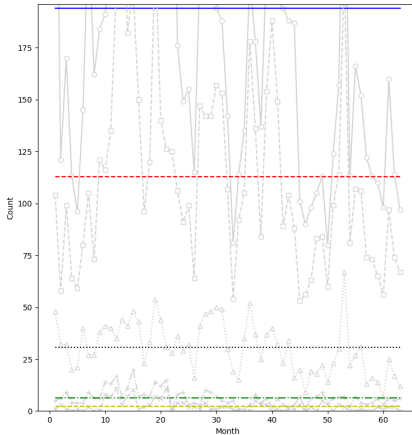
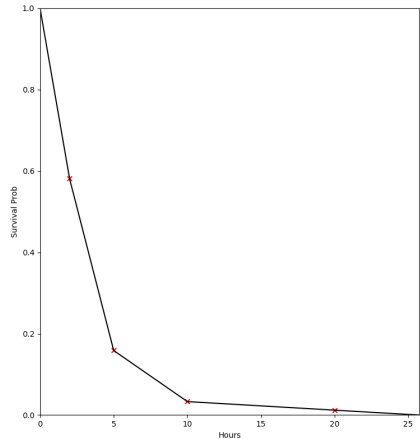
Rough method for severities



Rough method for severities



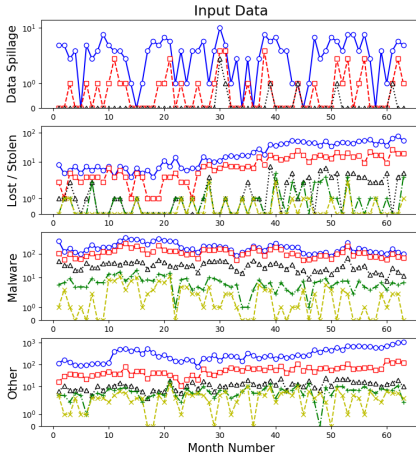
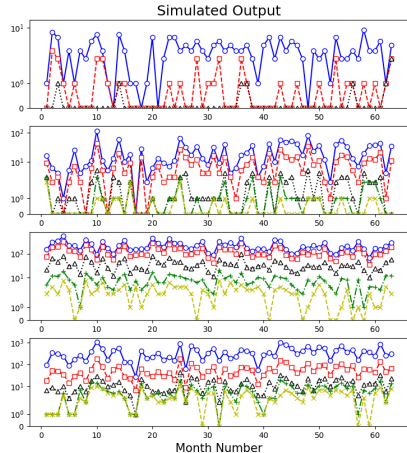
Rough method for severities



How does the model look?



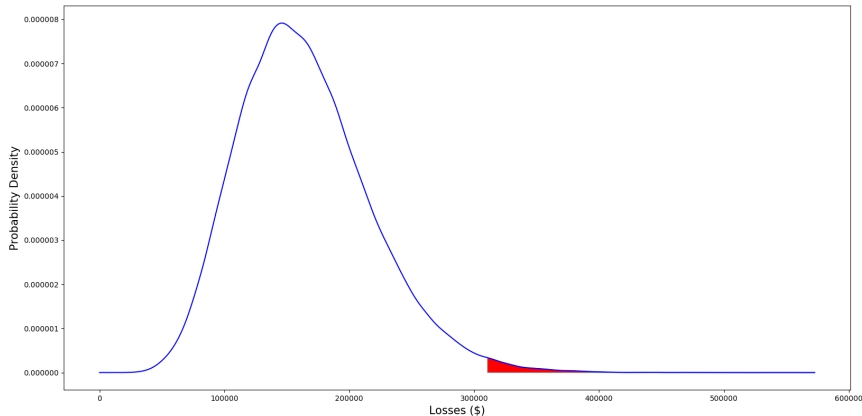
How does the model look?



Output to give to Angry Boss



Output to give to Angry Boss





“Number, shnumber! What do we do?!”



Simple model for mitigations



Simple model for mitigations

- M possible mitigations, max costs ω_m .



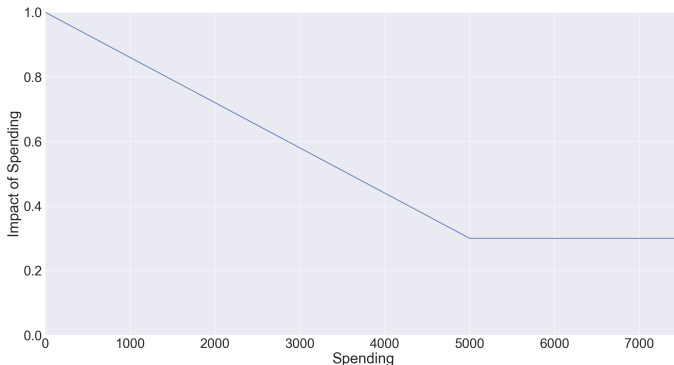
Simple model for mitigations



- M possible mitigations, max costs ω_m .
- Overall budget z , individual spendings z_m (sum to less than z).

Simple model for mitigations

- M possible mitigations, max costs ω_m .
- Overall budget z , individual spendings z_m (sum to less than z).
- Piecewise linear function f :



Simple model for mitigations



Simple model for mitigations



- Probability of successful attack of type j , given spendings on mitigations:

$$q_j(z_1, \dots, z_M) =$$

Simple model for mitigations



- Probability of successful attack of type j , given spendings on mitigations:

$$q_j(z_1, \dots, z_M) = q_j^0$$

Simple model for mitigations



- Probability of successful attack of type j , given spendings on mitigations:

$$q_j(z_1, \dots, z_M) = q_j^0 \prod_{m=1}^M f(z_m; \theta_{mj})$$

Simple model for mitigations



- Probability of successful attack of type j , given spendings on mitigations:

$$q_j(z_1, \dots, z_M) = \max \left(q_j^*, q_j^0 \prod_{m=1}^M f(z_m; \theta_{mj}) \right)$$

Simple model for mitigations



- Probability of successful attack of type j , given spendings on mitigations:

$$q_j(z_1, \dots, z_M) = \max \left(q_j^*, q_j^0 \prod_{m=1}^M f(z_m; \theta_{mj}) \right)$$

- Severity given successful attack of type j , given spendings on mitigations:

$$X_i^j(z_1, \dots, z_M) = \max \left(y_j^*, Y_i^j \prod_{m=1}^M f(z_m; \phi_{mj}) + \sum_{m=1}^M f(z_m; \psi_{mj}) \right)$$

Where to get values?



Where to get values?

Mitigation strategies summary



Relative security effectiveness rating	Mitigation strategy	Potential user resistance	Upfront cost (staff, equipment, technical complexity)	Ongoing maintenance cost (mainly staff)
Mitigation strategies to prevent malware delivery and execution				
Essential	Application whitelisting of approved/trusted programs to prevent execution of unapproved/malicious programs including .exe, DLL, scripts (e.g. Windows Script Host, PowerShell and HTA) and installers.	Medium	High	Medium
Essential	Patch applications e.g. Flash, web browsers, Microsoft Office, Java and PDF viewers. Patch/mitigate computers with 'extreme risk' vulnerabilities within 48 hours. Use the latest version of applications.	Low	High	High
Essential	Configure Microsoft Office macro settings to block macros from the Internet, and only allow vetted macros either in 'trusted locations' with limited write access or digitally signed with a trusted certificate.	Medium	Medium	Medium
Essential	User application hardening. Configure web browsers to block Flash (ideally uninstall it), ads and Java on the Internet. Disable unneeded features in Microsoft Office (e.g. OLE), web browsers and PDF viewers.	Medium	Medium	Medium
Excellent	Automated dynamic analysis of email and web content run in a sandbox , blocked if suspicious behaviour is identified e.g. network traffic, new or modified files, or other system configuration changes.	Low	High	Medium

Five mitigations



Five mitigations



- Application Whitelisting; effective for Malware, costs 200,000 per year.

Five mitigations



- Application Whitelisting; effective for Malware, costs 200,000 per year.
- Patch Applications; effective for “Other”, costs 200,000 per year.

Five mitigations



- Application Whitelisting; effective for Malware, costs 200,000 per year.
- Patch Applications; effective for “Other”, costs 200,000 per year.
- Continuous Incident Detection; effective for Data Spillage and Lost / Stolen, costs 300,000 per year.

Five mitigations



- Application Whitelisting; effective for Malware, costs 200,000 per year.
- Patch Applications; effective for “Other”, costs 200,000 per year.
- Continuous Incident Detection; effective for Data Spillage and Lost / Stolen, costs 300,000 per year.
- Antivirus Software; effective for Malware, costs 150,000 per year

Five mitigations



- Application Whitelisting; effective for Malware, costs 200,000 per year.
- Patch Applications; effective for “Other”, costs 200,000 per year.
- Continuous Incident Detection; effective for Data Spillage and Lost / Stolen, costs 300,000 per year.
- Antivirus Software; effective for Malware, costs 150,000 per year
- TLS Encryption; effective for “Other”, costs 150,000 per year.

Five mitigations

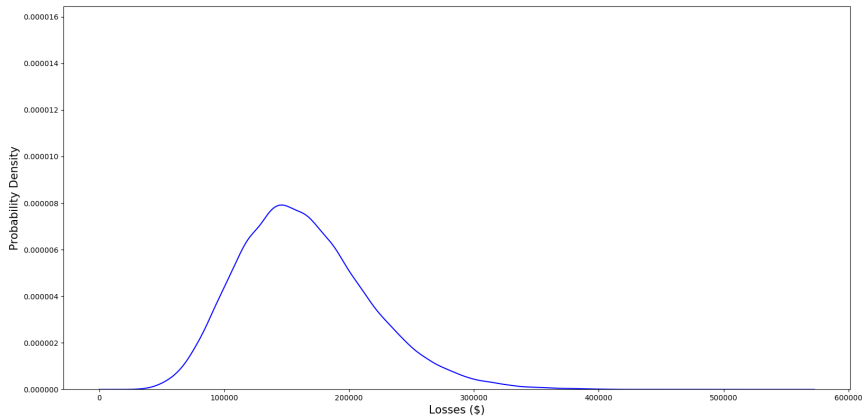


- Application Whitelisting; effective for Malware, costs 200,000 per year.
- Patch Applications; effective for “Other”, costs 200,000 per year.
- Continuous Incident Detection; effective for Data Spillage and Lost / Stolen, costs 300,000 per year.
- Antivirus Software; effective for Malware, costs 150,000 per year
- TLS Encryption; effective for “Other”, costs 150,000 per year.
- Sneaky assumptions!

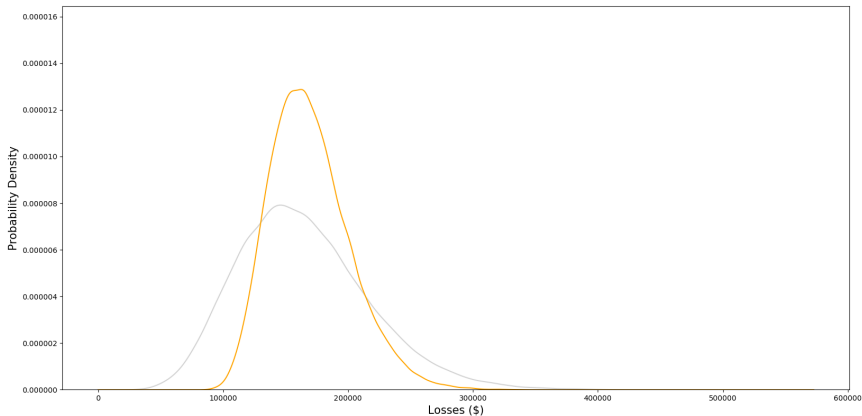
No Mitigations Losses



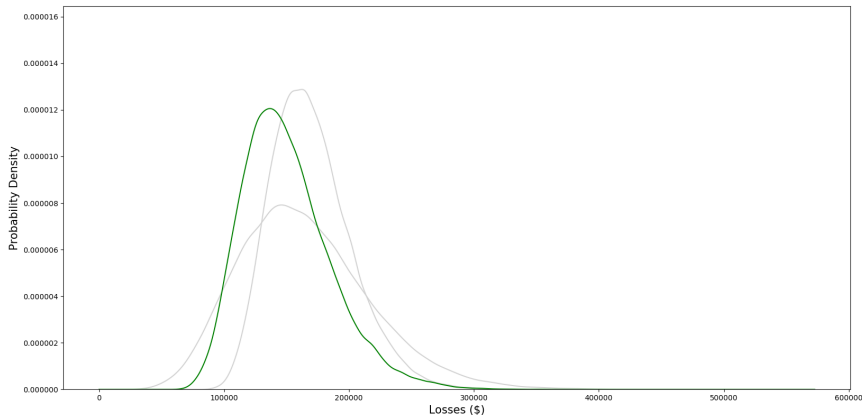
No Mitigations Losses



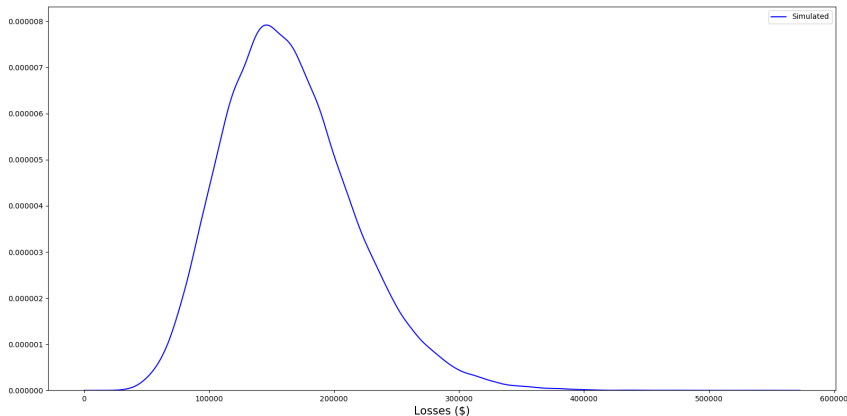
Equal Spendings



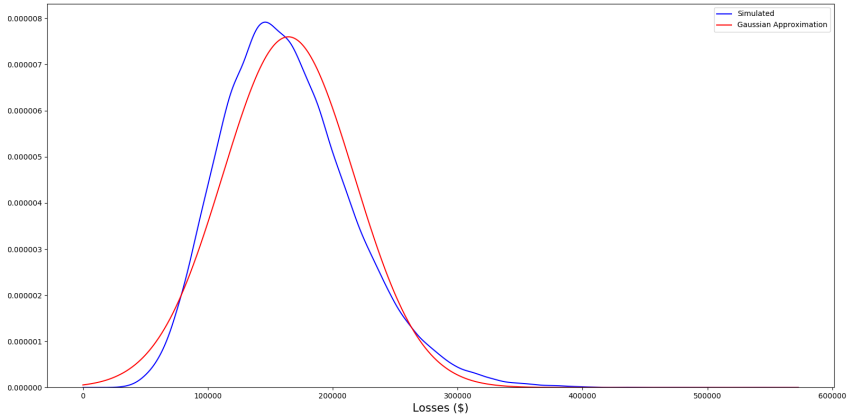
Optimising for the Mean



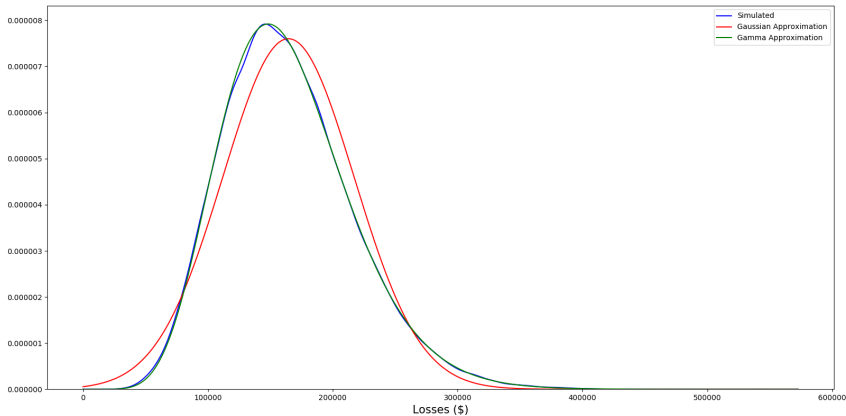
Gaussian Approximation



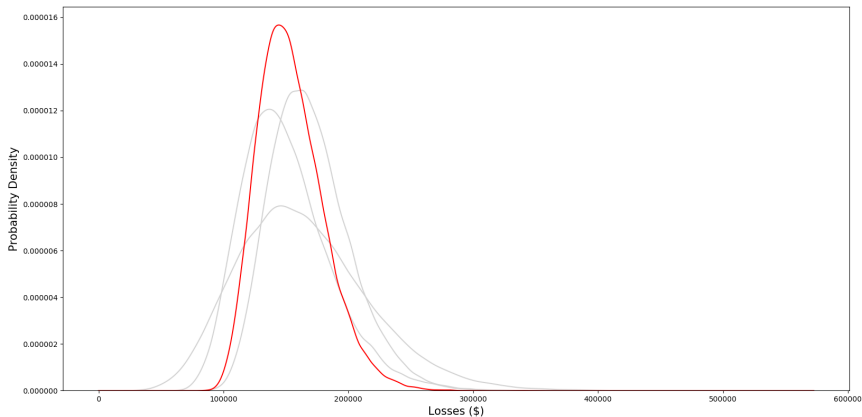
Gaussian Approximation



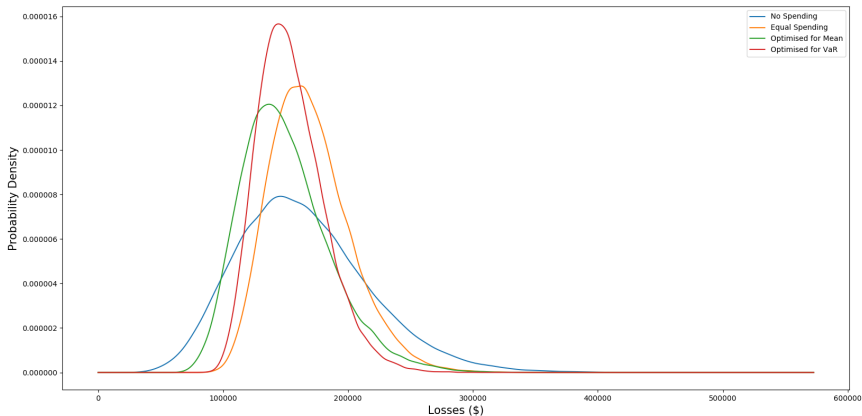
Shifted Gamma Approximation



Optimising for the VaR



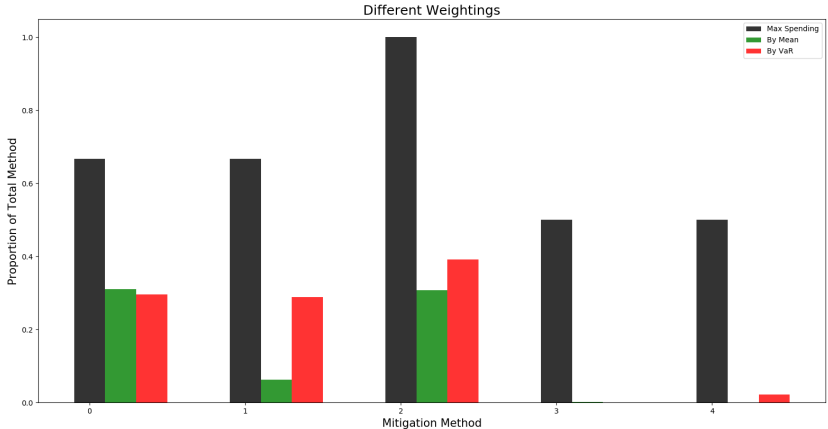
Different Strategies



Weights not quite right



Weights not quite right



Do the ρ_{ij} actually do much?



Do the ρ_{ij} actually do much?



Consider a measure of overall dependence as

$$\text{dependence} = \frac{\sum_{i \neq j} \rho_{ij}}{J^2 - J}$$

Do the ρ_{ij} actually do much?



Consider a measure of overall dependence as

$$\text{dependence} = \frac{\sum_{i \neq j} \rho_{ij}}{J^2 - J}$$

and a measure of overall impact of the correlations as

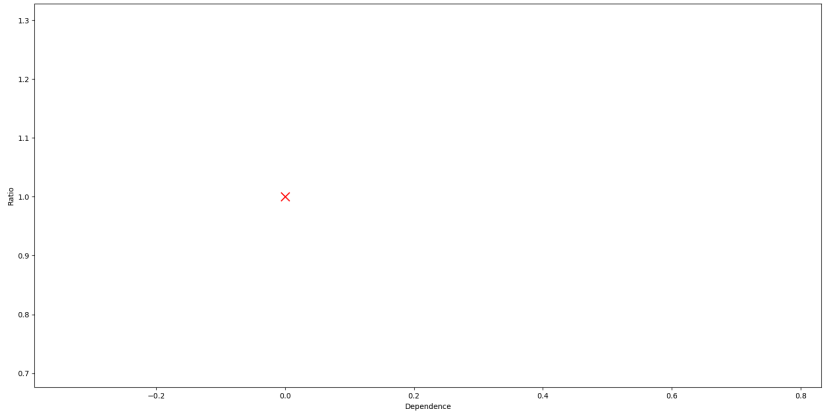
$$\frac{\text{VaR}(\text{dependence})}{\text{VaR}(0)}$$

Do the ρ_{ij} actually do much?



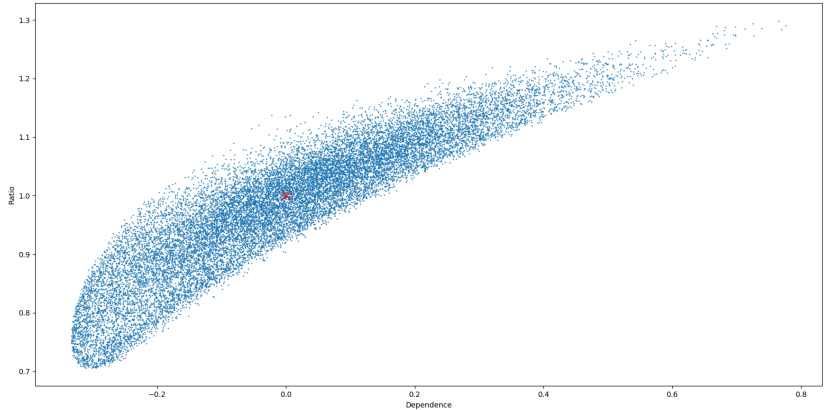
Do the ρ_{ij} actually do much?

Nope!

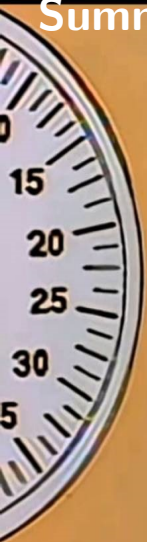


Do the ρ_{ij} actually do much?

But they could...

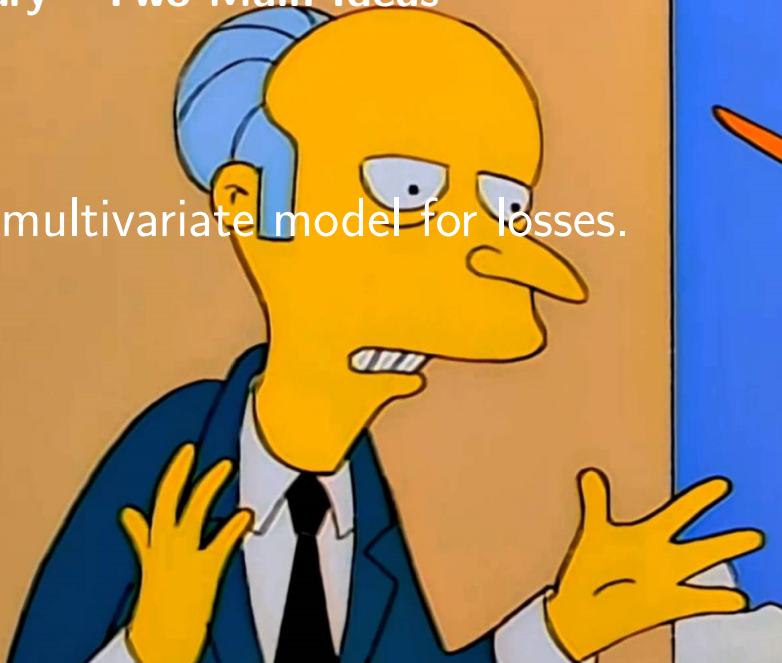


Summary - Two Main Ideas



Summary - Two Main Ideas

- A multivariate model for losses.



Summary - Two Main Ideas

- A multivariate model for losses.
- A framework to suggest mitigations.



Idea I - Multivariate Model



Idea 1 - Multivariate Model



- Use a Gaussian copula:

Idea I - Multivariate Model



$$L(t) = \sum_{j=1}^J \sum_{i=1}^{N_j(t)} x_i^j$$

- Use a Gaussian copula:

Idea I - Multivariate Model



$$L(t) = \sum_{j=1}^J \sum_{i=1}^{N_j(t)} x_i^j$$

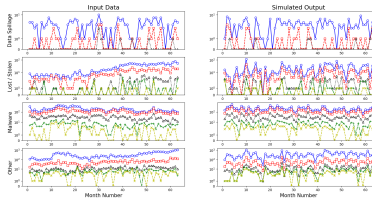
- Use a Gaussian copula:

- Works well enough:

Idea I - Multivariate Model

$$L(t) = \sum_{j=1}^J \sum_{i=1}^{N_j(t)} X_i^j$$

- Use a Gaussian copula:

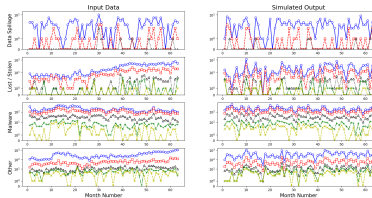


- Works well enough:

Idea I - Multivariate Model

$$L(t) = \sum_{j=1}^J \sum_{i=1}^{N_j(t)} X_i^j$$

- Use a Gaussian copula:

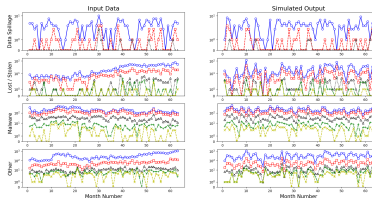


- Works well enough:
- Output a distribution:

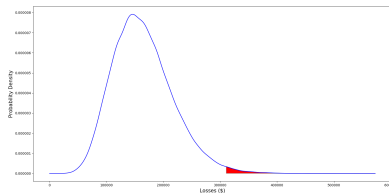
Idea I - Multivariate Model

$$L(t) = \sum_{j=1}^J \sum_{i=1}^{N_j(t)} X_i^j$$

- Use a Gaussian copula:



- Works well enough:



- Output a distribution:

Idea II - A framework for mitigations



Idea II - A framework for mitigations



- Gordon-Loeb:

Idea II - A framework for mitigations



- Gordon-Loeb: (thinning)
$$q_j(z_1, \dots, z_M) = \max \left(q_j^*, q_j^0 \prod_{m=1}^M f(z_m; \theta_{mj}) \right)$$

Idea II - A framework for mitigations



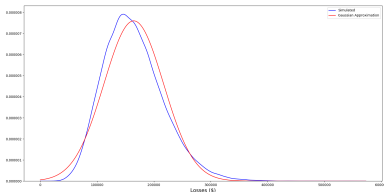
$$q_j(z_1, \dots, z_M) = \max \left(q_j^*, q_j^0 \prod_{m=1}^M f(z_m; \theta_{mj}) \right)$$

- Gordon-Loeb: (thinning)
- Use a closed form approximation ... :

Idea II - A framework for mitigations

$$q_j(z_1, \dots, z_M) = \max \left(q_j^*, q_j^0 \prod_{m=1}^M f(z_m; \theta_{mj}) \right)$$

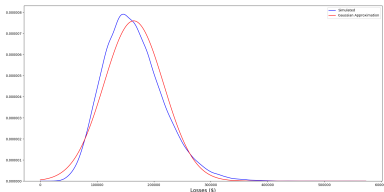
- Gordon-Loeb: (thinning)
- Use a closed form approximation ... :



Idea II - A framework for mitigations

$$q_j(z_1, \dots, z_M) = \max \left(q_j^*, q_j^0 \prod_{m=1}^M f(z_m; \theta_{mj}) \right)$$

- Gordon-Loeb: (thinning)
- Use a closed form approximation ... :

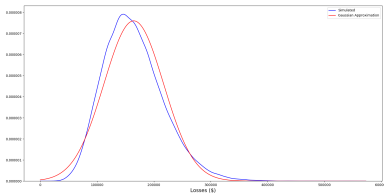
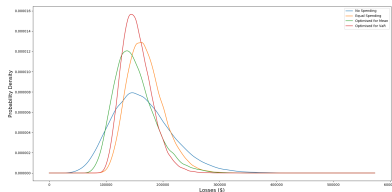


- ... to optimise:

Idea II - A framework for mitigations

$$q_j(z_1, \dots, z_M) = \max \left(q_j^*, q_j^0 \prod_{m=1}^M f(z_m; \theta_{mj}) \right)$$

- Gordon-Loeb: (thinning)
- Use a closed form approximation ... :



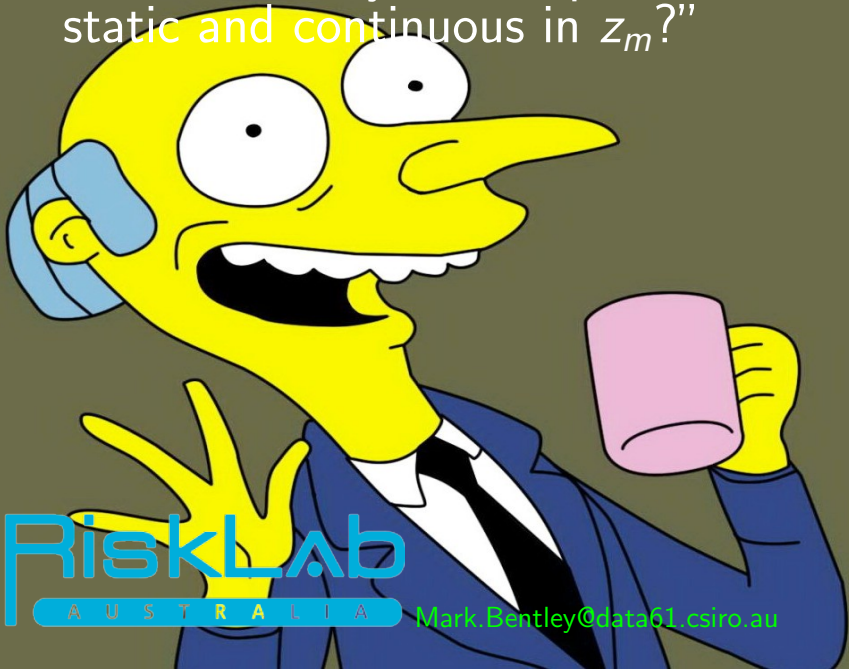
- ... to optimise:



- “Nice, but why is the optimisation static and continuous in z_m ?”



- “Nice, but why is the optimisation static and continuous in z_m ?”

A cartoon illustration of a yellow-skinned man with a large, pointed nose, wide eyes, and a large open-mouthed smile. He is wearing a blue suit jacket, a white shirt, and a black tie. He has a blue bandage on his left ear and is holding a pink rectangular object in his right hand. His left hand is raised with fingers spread.

RiskLab
A U S T R A L I A

Mark.Bentley@data61.csiro.au