# Seminar on Reinsurance

June 5-6, 2023

Westin Philadelphia

Philadelphia, PA

# SIMULATION UNCERTAINTY

## How Reliable are Our Tail Statistics?

June 2023

Dean Marcus, FCAS CERA | SVP Actuary | New York

# Agenda

# 1. Background and Motivation

# Background and Motivation

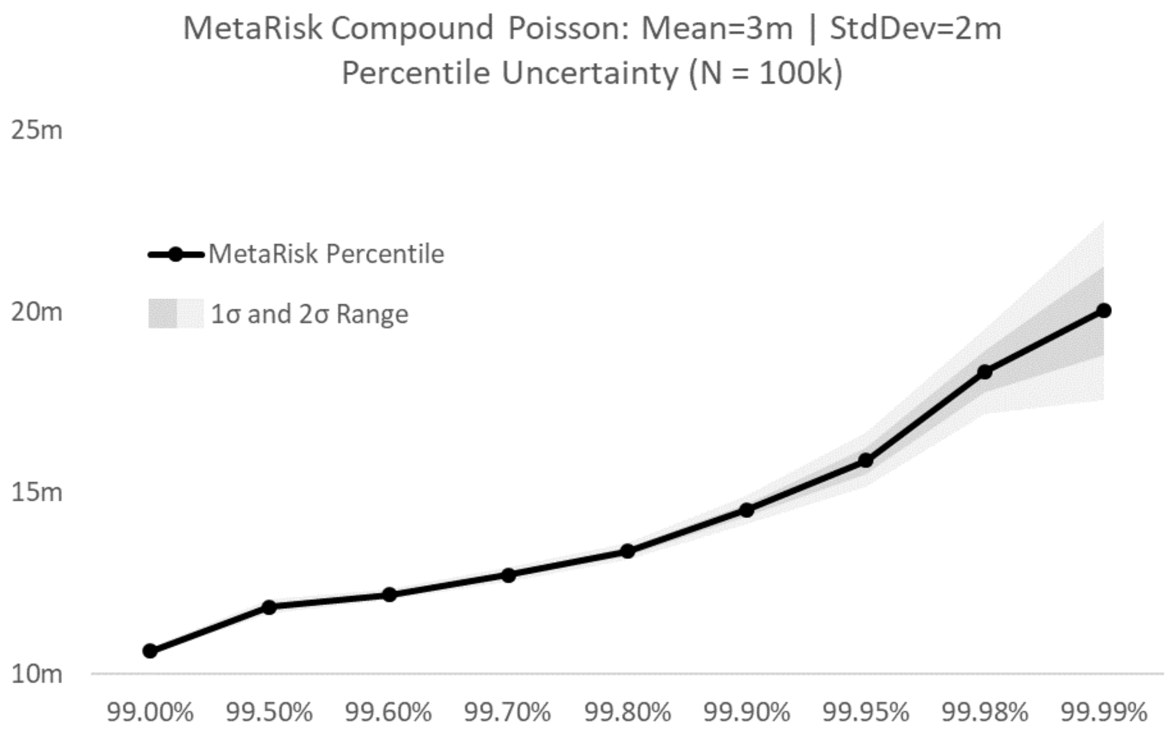
## Two Questions

1. **What is the statistical nature of our simulated percentiles?** What can I say about the *error bands around* the percentiles, and are the error bands *themselves* a bit stretchy/blurry with uncertainty?

2. **How many realizations should we run?** The answer will depend on the context, but our decisions should be well-informed by empirical evidence and solid theory, including the answer to question 1

Addressing these two questions will help guide practical decisions, and ensure sound advice to brokers and clients

| | Gross Loss and ALAE | | |
|---|---|---|---|
| Probability | VaR | Uncertainty | TVaR |
| 95.00% | $7,047,493 | 0.42% | $9,242,559 |
| 98.00% | $9,243,777 | 0.57% | $11,132,645 |
| 99.00% | $10,625,354 | 0.43% | $12,391,280 |
| 99.50% | $11,853,589 | 0.85% | $13,635,187 |
| 99.60% | $12,188,010 | 0.57% | $14,037,238 |
| 99.80% | $13,376,645 | 0.87% | $15,344,299 |
| 99.90% | $14,522,656 | 1.30% | $16,824,069 |
| 99.99% | $20,026,814 | 6.19% | $22,005,882 |



# Background and Motivation

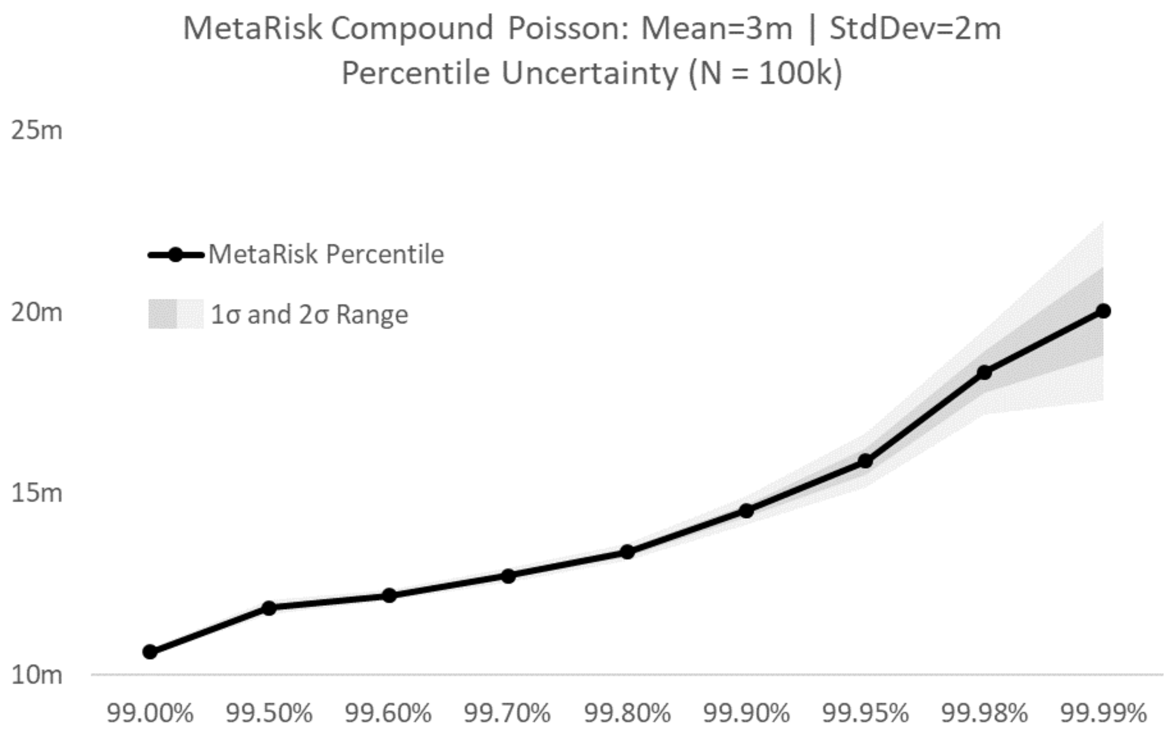
## Two Questions

1. **What is the statistical nature of our simulated percentiles?** What can I say about the *error bands around* the percentiles, and are the error bands *themselves* a bit stretchy/blurry with uncertainty?

2. **How many realizations should we run?** The answer will depend on the context, but our decisions should be well-informed by empirical evidence and solid theory, including the answer to question 1

Addressing these two questions will help guide practical decisions, and ensure sound advice to brokers and clients

| | Gross Loss and ALAE | | |
|---|---|---|---|
| Probability | VaR | Uncertainty | TVaR |
| 95.00% | $7,047,493 | 0.42% | $9,242,559 |
| 98.00% | $9,243,777 | 0.57% | $11,132,645 |
| 99.00% | $10,625,354 | 0.43% | $12,391,280 |
| 99.50% | $11,853,589 | 0.85% | $13,635,187 |
| 99.60% | $12,188,010 | 0.57% | $14,037,238 |
| 99.80% | $13,376,645 | 0.87% | $15,344,299 |
| 99.90% | $14,522,656 | 1.30% | $16,824,069 |
| 99.99% | $20,026,814 | 6.19% | $22,005,882 |



# Background and Motivation

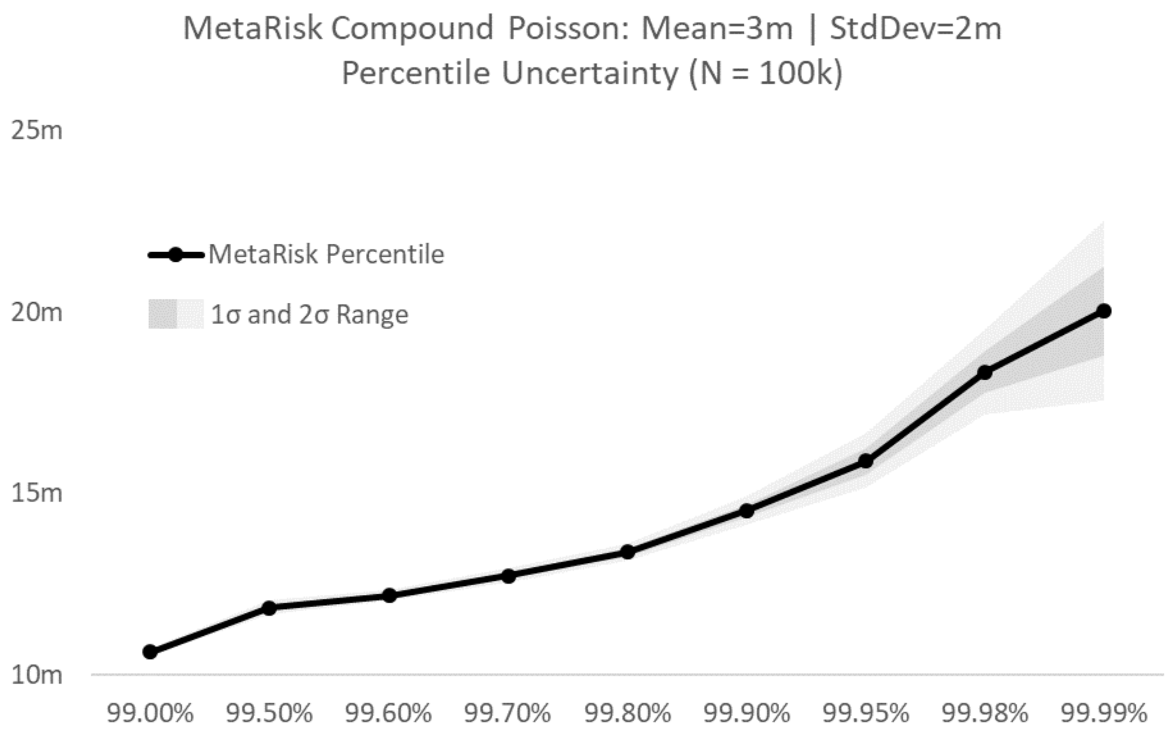
## Two Questions

1. **What is the statistical nature of our simulated percentiles?** What can I say about the *error bands around* the percentiles, and are the error bands *themselves* a bit stretchy/blurry with uncertainty?

2. **How many realizations should we run?** The answer will depend on the context, but our decisions should be well-informed by empirical evidence and solid theory, including the answer to question 1

Addressing these two questions will help guide practical decisions, and ensure sound advice to brokers and clients

| | Gross Loss and ALAE | | |
|---|---|---|---|
| Probability | VaR | Uncertainty | TVaR |
| 95.00% | $7,047,493 | 0.42% | $9,242,559 |
| 98.00% | $9,243,777 | 0.57% | $11,132,645 |
| 99.00% | $10,625,354 | 0.43% | $12,391,280 |
| 99.50% | $11,853,589 | 0.85% | $13,635,187 |
| 99.60% | $12,188,010 | 0.57% | $14,037,238 |
| 99.80% | $13,376,645 | 0.87% | $15,344,299 |
| 99.90% | $14,522,656 | 1.30% | $16,824,069 |
| 99.99% | $20,026,814 | 6.19% | $22,005,882 |



# Background and Motivation

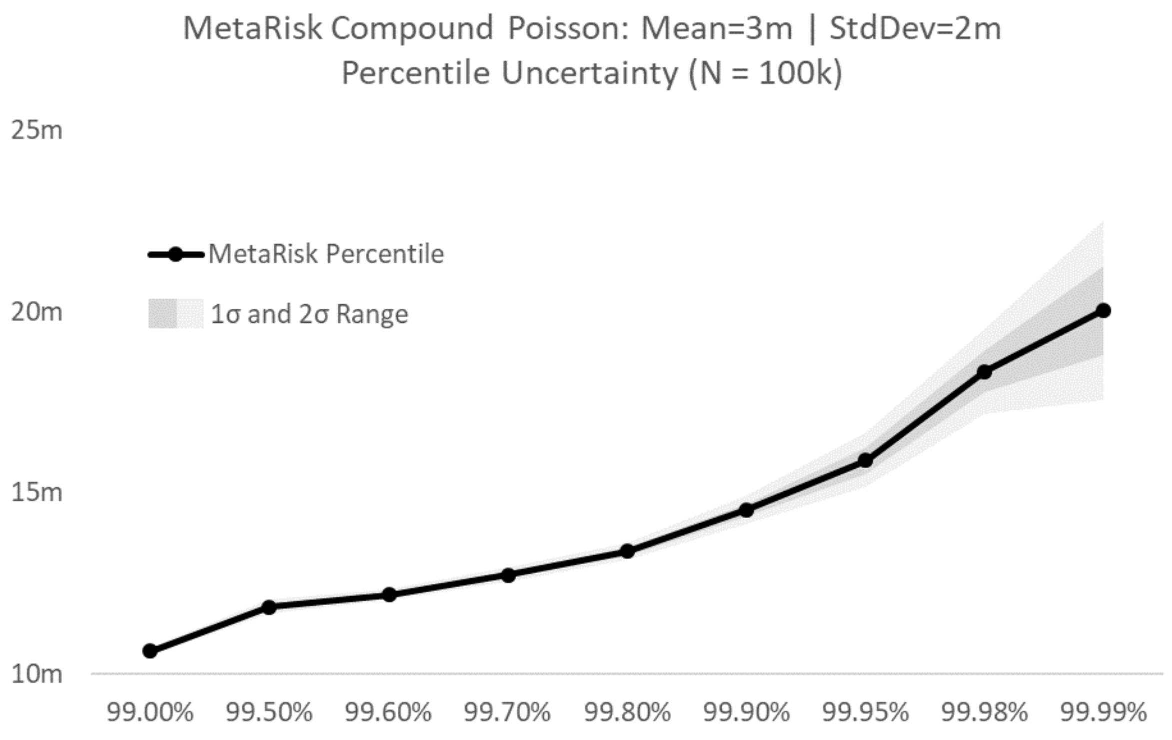
## Two Questions

1. **What is the statistical nature of our simulated percentiles?** What can I say about the *error bands around* the percentiles, and are the error bands *themselves* a bit stretchy/blurry with uncertainty?

2. **How many realizations should we run?** The answer will depend on the context, but our decisions should be well-informed by empirical evidence and solid theory, including the answer to question 1

Addressing these two questions will help guide practical decisions, and ensure sound advice to brokers and clients

| | Gross Loss and ALAE | | |
|---|---|---|---|
| Probability | VaR | Uncertainty | TVaR |
| 95.00% | $7,047,493 | 0.42% | $9,242,559 |
| 98.00% | $9,243,777 | 0.57% | $11,132,645 |
| 99.00% | $10,625,354 | 0.43% | $12,391,280 |
| 99.50% | $11,853,589 | 0.85% | $13,635,187 |
| 99.60% | $12,188,010 | 0.57% | $14,037,238 |
| 99.80% | $13,376,645 | 0.87% | $15,344,299 |
| 99.90% | $14,522,656 | 1.30% | $16,824,069 |
| 99.99% | $20,026,814 | 6.19% | $22,005,882 |



# Background and Motivation

## Sample Mean and CV

If we re-simulate any given model/metric using a different random seed, the key results will change… but they converge with high samples:

- **Sample mean** is well-understood and close to the theoretical mean – its CV is $\frac{True\ CV}{\sqrt{\#Samples}}$, so, e.g., re-simulating 1m realizations on a distribution that has CV 250% will typically produce a new mean within ≈0.25% of the previous one
- **Sample standard deviation** is well-understood but converges relatively slowly to the theoretical standard deviation
- **Lognormal example below**: 10k samples produces reliable estimates for the mean, but unreliable estimates for the CV, whereas 500k samples converges well

# Background and Motivation

## Sample Mean and CV

If we re-simulate any given model/metric using a different random seed, the key results will change… but they converge with high samples:

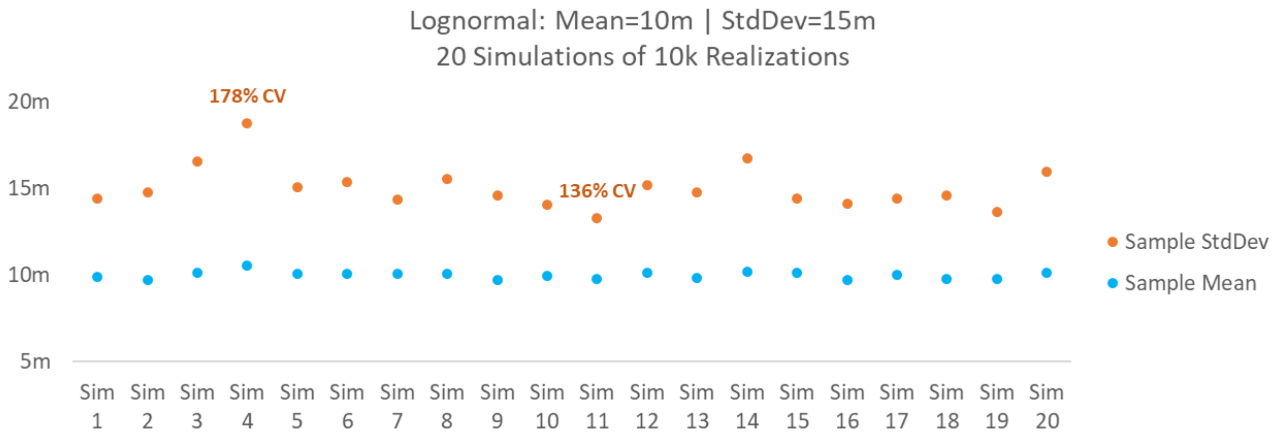
- **Sample mean** is well-understood and close to the theoretical mean – its CV is $\frac{True\ CV}{\sqrt{\#Samples}}$, so, e.g., re-simulating 1m realizations on a distribution that has CV 250% will typically produce a new mean within ≈0.25% of the previous one
- **Sample standard deviation** is well-understood but converges relatively slowly to the theoretical standard deviation
- **Lognormal example below**: 10k samples produces reliable estimates for the mean, but unreliable estimates for the CV, whereas 500k samples converges well

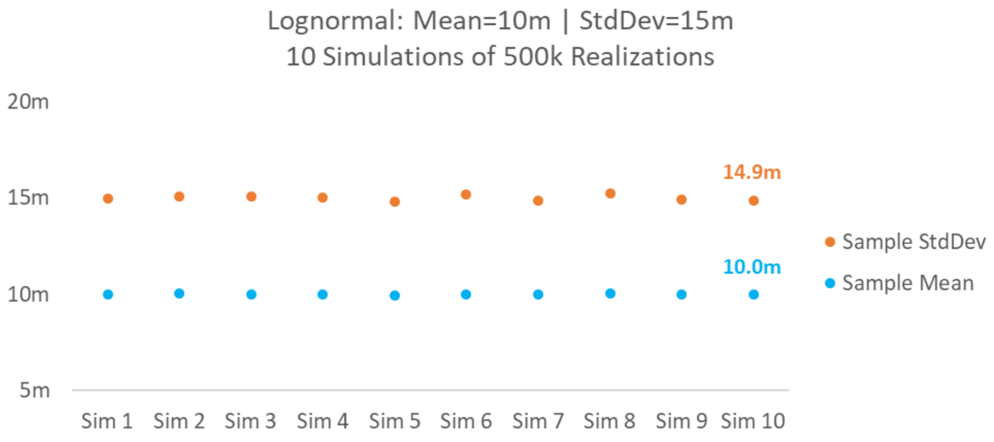
# Background and Motivation

## Sample Mean and CV

If we re-simulate any given model/metric using a different random seed, the key results will change… but they converge with high samples:

- **Sample mean** is well-understood and close to the theoretical mean – its CV is $\frac{True\ CV}{\sqrt{\#Samples}}$, so, e.g., re-simulating 1m realizations on a distribution that has CV 250% will typically produce a new mean within ≈0.25% of the previous one
- **Sample standard deviation** is well-understood but converges relatively slowly to the theoretical standard deviation
- **Lognormal example below**: 10k samples produces reliable estimates for the mean, but unreliable estimates for the CV, whereas 500k samples converges well



# Background and Motivation

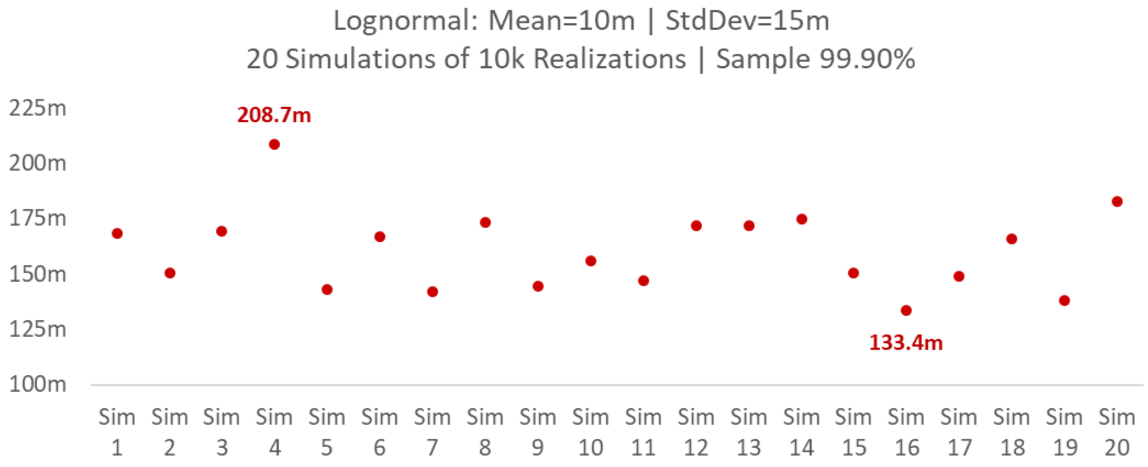
## Sample Percentiles

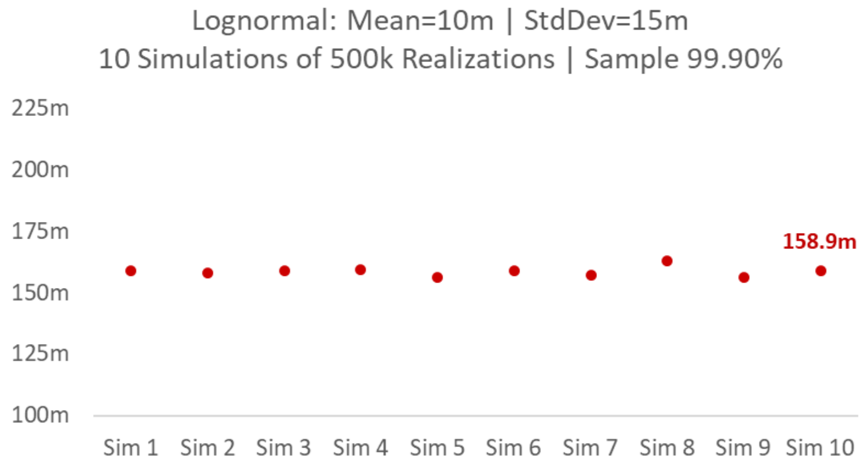
- **Sample percentiles aren't discussed as much and can vary wildly!** This will be our focus, with very interesting findings…
- **Lognormal example below**: 10k samples produces unreliable estimates for the tail stats, whereas 500k samples converges well

# Background and Motivation

## Sample Percentiles

- **Sample percentiles aren't discussed as much and can vary wildly!** This will be our focus, with very interesting findings…
- **Lognormal example below**: 10k samples produces unreliable estimates for the tail stats, whereas 500k samples converges well

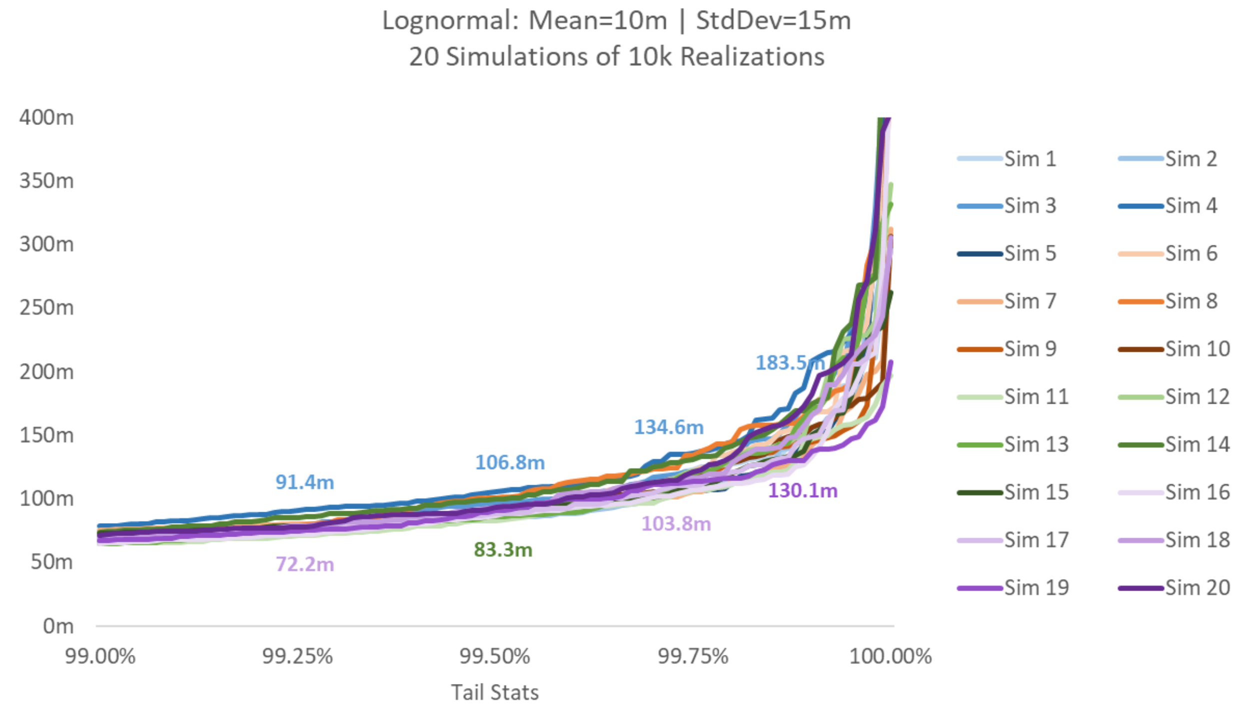


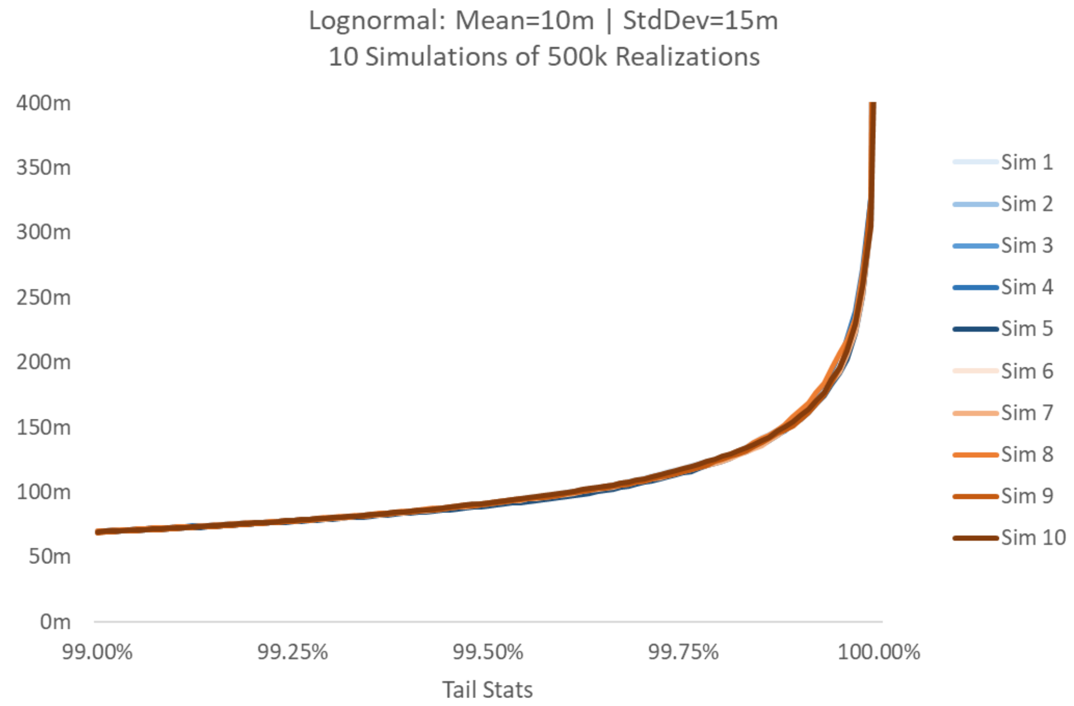



# Background and Motivation

## Sample Percentiles

- **Sample percentiles aren't discussed as much and can vary wildly!** This will be our focus, with very interesting findings…
- **Lognormal example below**: 10k samples produces unreliable estimates for the tail stats, whereas 500k samples converges well



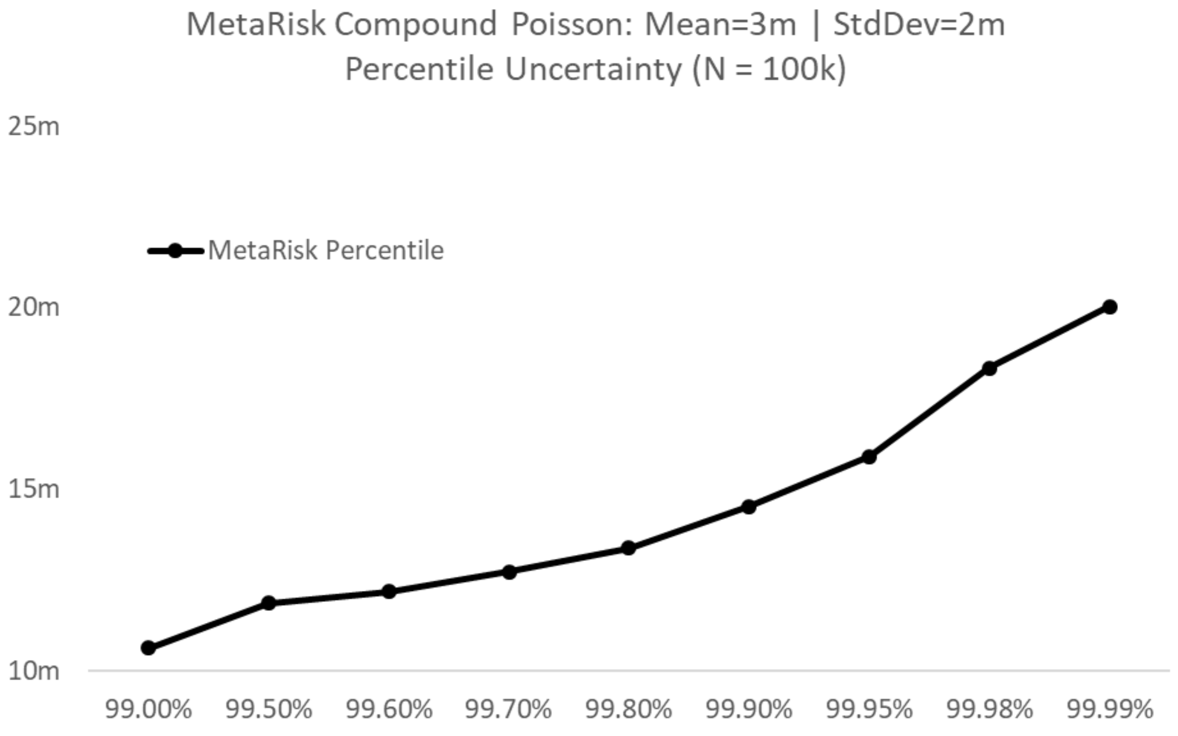



# Background and Motivation

## Uncertainty Frameworks

So we can think of simulation uncertainty as having two levels:

1. **First-Order Uncertainty**: Any simulated metric (random variable) like gross losses, net losses, ceded to a contract, etc. has percentiles – i.e., **a CDF**
2. **Second-Order Uncertainty**: The simulated percentiles *themselves* have inherent uncertainty and will change if we re-simulate – i.e., *our attempt at* **Error Bars** using the imperfect information from our simulation

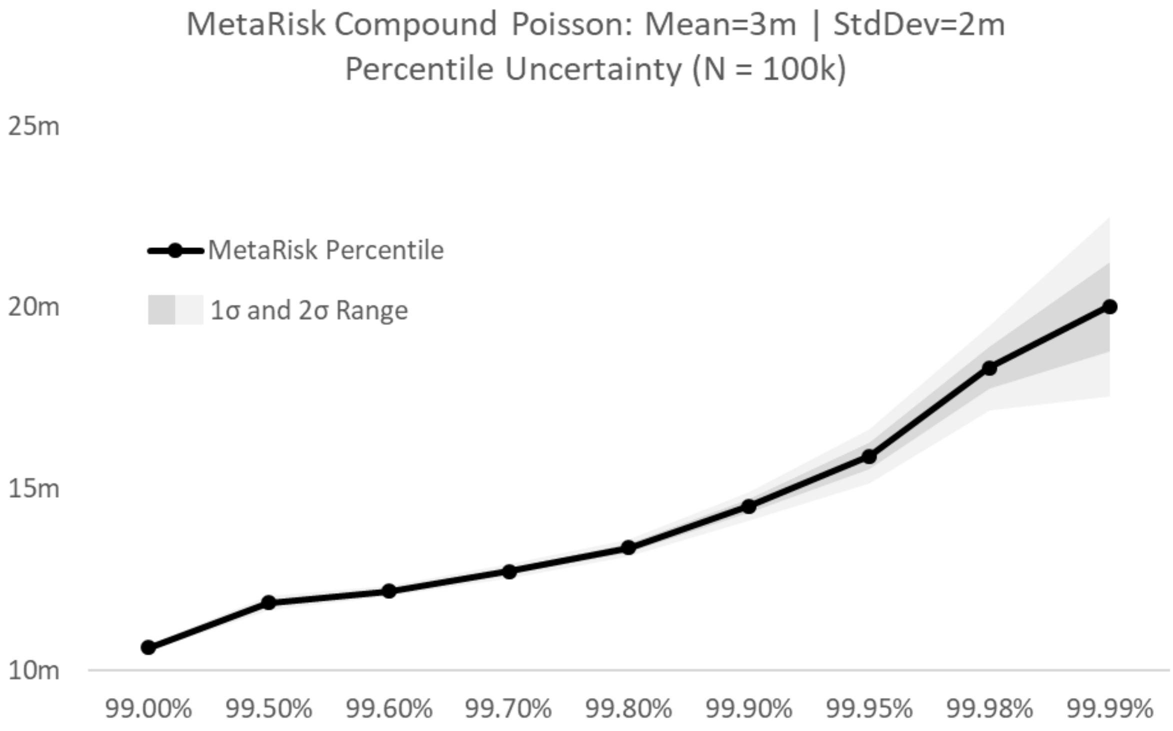


# Background and Motivation

## Uncertainty Frameworks

So we can think of simulation uncertainty as having two levels:

1. **First-Order Uncertainty**: Any simulated metric (random variable) like gross losses, net losses, ceded to a contract, etc. has percentiles – i.e., **a CDF**
2. **Second-Order Uncertainty**: The simulated percentiles *themselves* have inherent uncertainty and will change if we re-simulate – i.e., *our attempt at* **Error Bars** using the imperfect information from our simulation



# 2. Spoiler Alert: Key Takeaways

# Key Takeaways

There is a well-known formula to estimate the CV upon re-simulation, but it uses the *simulated* gradient of the CDF around the percentile

The estimated CV of any simulated percentile *itself* varies on re-simulation with CV $\frac{1}{\sqrt{O}}$ because the *simulated* gradients are noisy… Stay tuned for $O$!

Percentiles for RMS Cat XOL models have very high simulation uncertainty, whereas Quota Shares and Per Risk XOLs seem to converge quicker

If your cat reinsurance decision is driven by RMS tail statistics and isn't diversified by region/peril then you will often need > 2m realizations!

# Key Takeaways

There is a well-known formula to estimate the CV upon re-simulation, but it uses the *simulated* gradient of the CDF around the percentile

The estimated CV of any simulated percentile *itself* varies on re-simulation with CV $\frac{1}{\sqrt{O}}$ because the *simulated* gradients are noisy… Stay tuned for $O$!

Percentiles for RMS Cat XOL models have very high simulation uncertainty, whereas Quota Shares and Per Risk XOLs seem to converge quicker

If your cat reinsurance decision is driven by RMS tail statistics and isn't diversified by region/peril then you will often need > 2m realizations!

# Key Takeaways

There is a well-known formula to estimate the CV upon re-simulation, but it uses the *simulated* gradient of the CDF around the percentile

The estimated CV of any simulated percentile *itself* varies on re-simulation with CV $\frac{1}{\sqrt{O}}$ because the *simulated* gradients are noisy… Stay tuned for $O$!

Percentiles for RMS Cat XOL models have very high simulation uncertainty, whereas Quota Shares and Per Risk XOLs seem to converge quicker

If your cat reinsurance decision is driven by RMS tail statistics and isn't diversified by region/peril then you will often need > 2m realizations!

# Key Takeaways

There is a well-known formula to estimate the CV upon re-simulation, but it uses the *simulated* gradient of the CDF around the percentile

The estimated CV of any simulated percentile *itself* varies on re-simulation with CV $\frac{1}{\sqrt{O}}$ because the *simulated* gradients are noisy… Stay tuned for $O$!

Percentiles for RMS Cat XOL models have very high simulation uncertainty, whereas Quota Shares and Per Risk XOLs seem to converge quicker

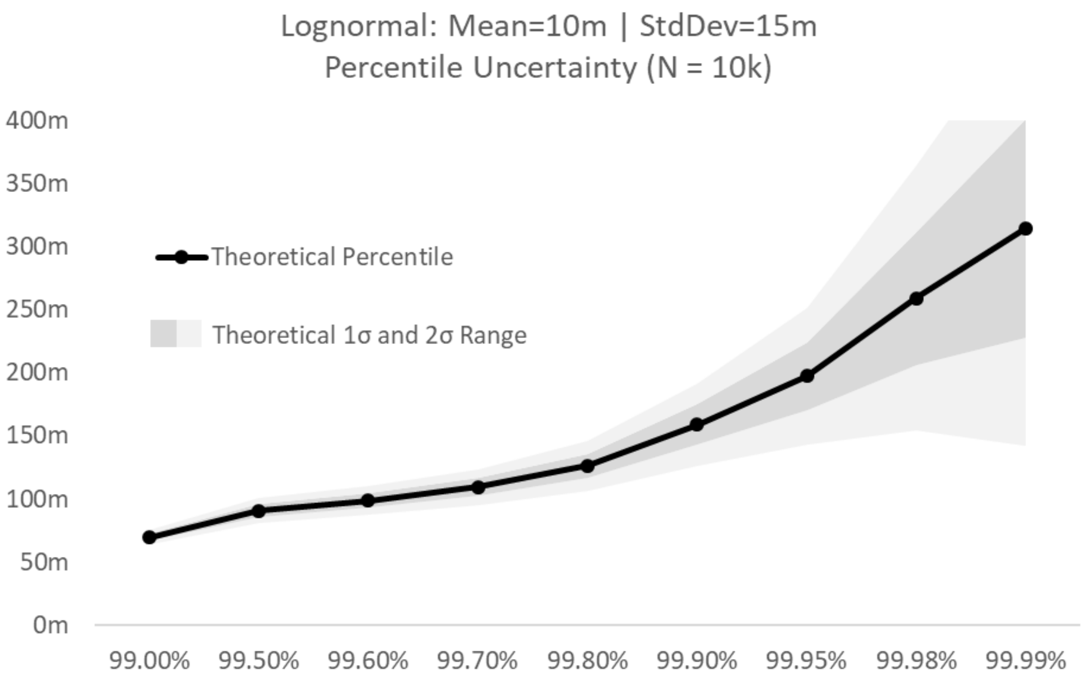
If your cat reinsurance decision is driven by RMS tail statistics and isn't diversified by region/peril then you will often need > 2m realizations!

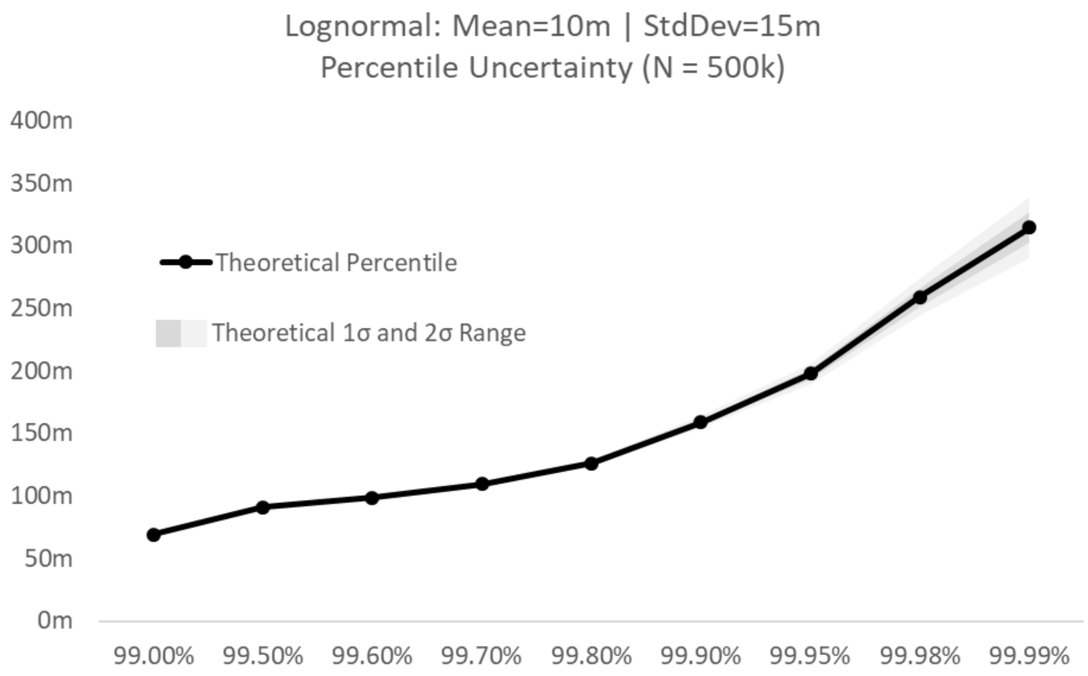
# 3. What’s the Deal with Second-Order Uncertainty?

# Second-Order Uncertainty

## Volatility of Percentile Estimates: Known Distributions

- **There is a well-known asymptotic formula**: $\sigma \to \sqrt{\frac{p(1-p)}{samples}} \times F^{-1'}(p)$, where *F(x)* is the CDF of the underlying distribution, so $F^{-1}$ is the graph below and $F^{-1'}$ is just the gradient of the graph below at each percentile
- **For high percentiles**, our estimates are more reliable as the samples increase or as the gradient decreases
- **Lognormal example below**: Theory matches our experiment on Slide 13 above!

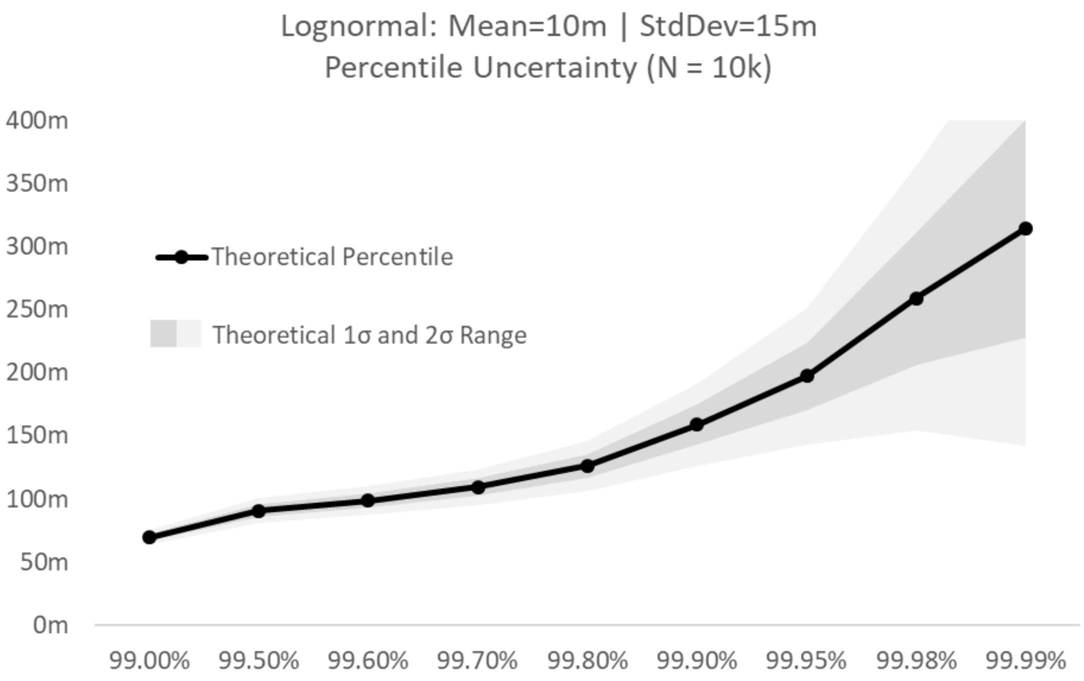


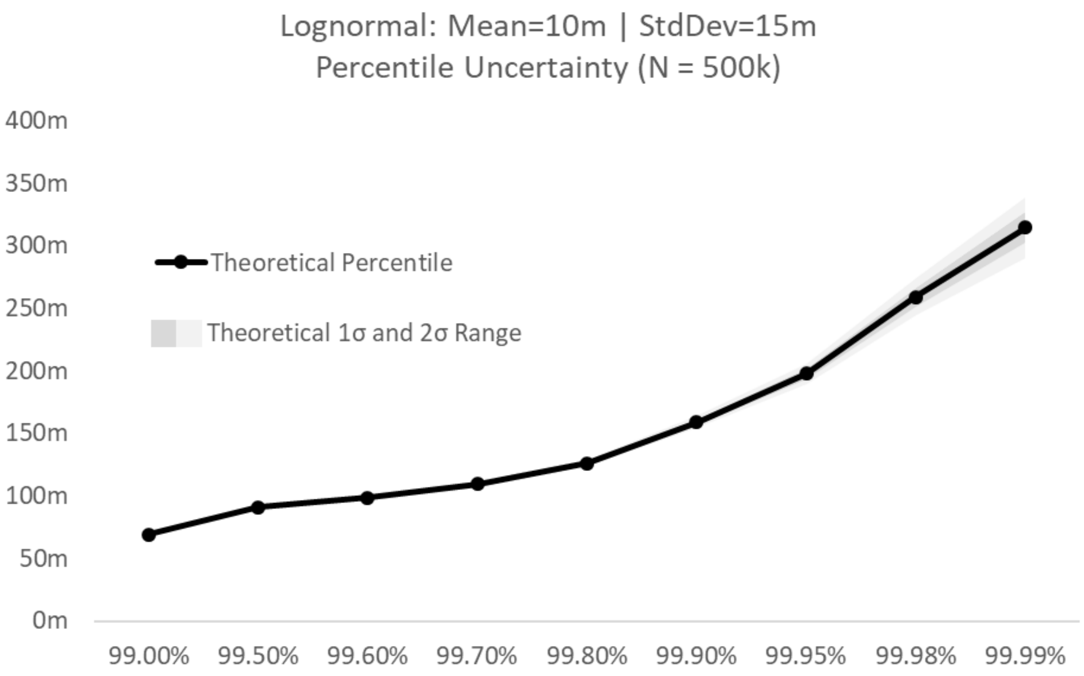



# Second-Order Uncertainty

## Volatility of Percentile Estimates: Known Distributions

- **There is a well-known asymptotic formula**: $\sigma \to \sqrt{\frac{p(1-p)}{samples}} \times F^{-1'}(p)$, where *F(x)* is the CDF of the underlying distribution, so $F^{-1}$ is the graph below and $F^{-1'}$ is just the gradient of the graph below at each percentile
- **For high percentiles**, our estimates are more reliable as the samples increase or as the gradient decreases
- **Lognormal example below**: Theory matches our experiment on Slide 13 above!

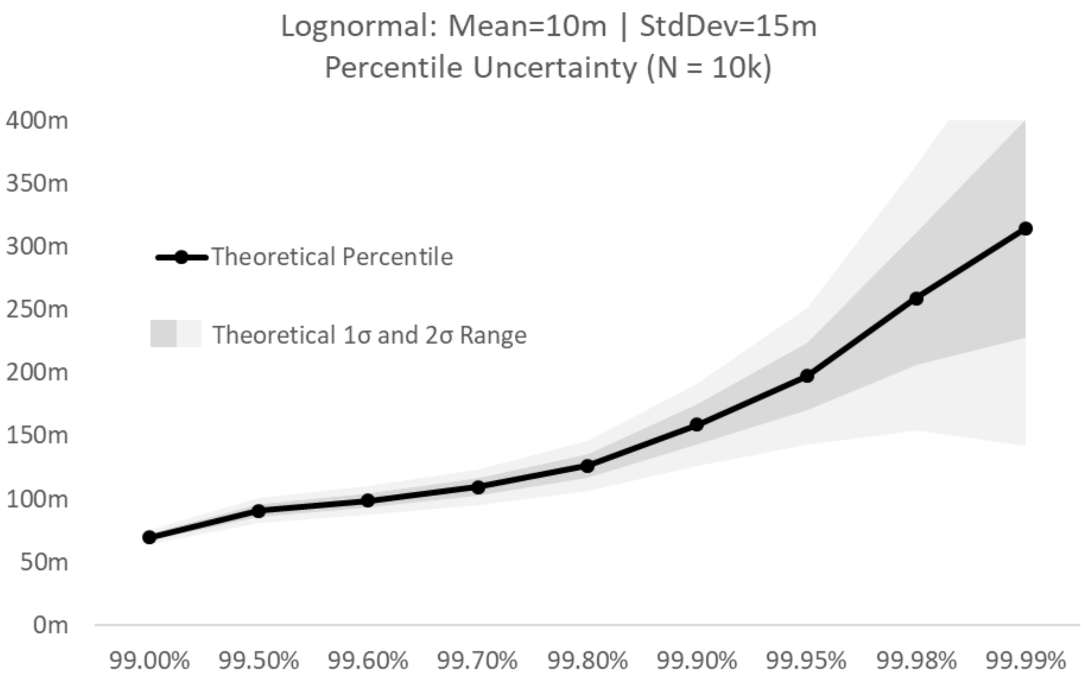


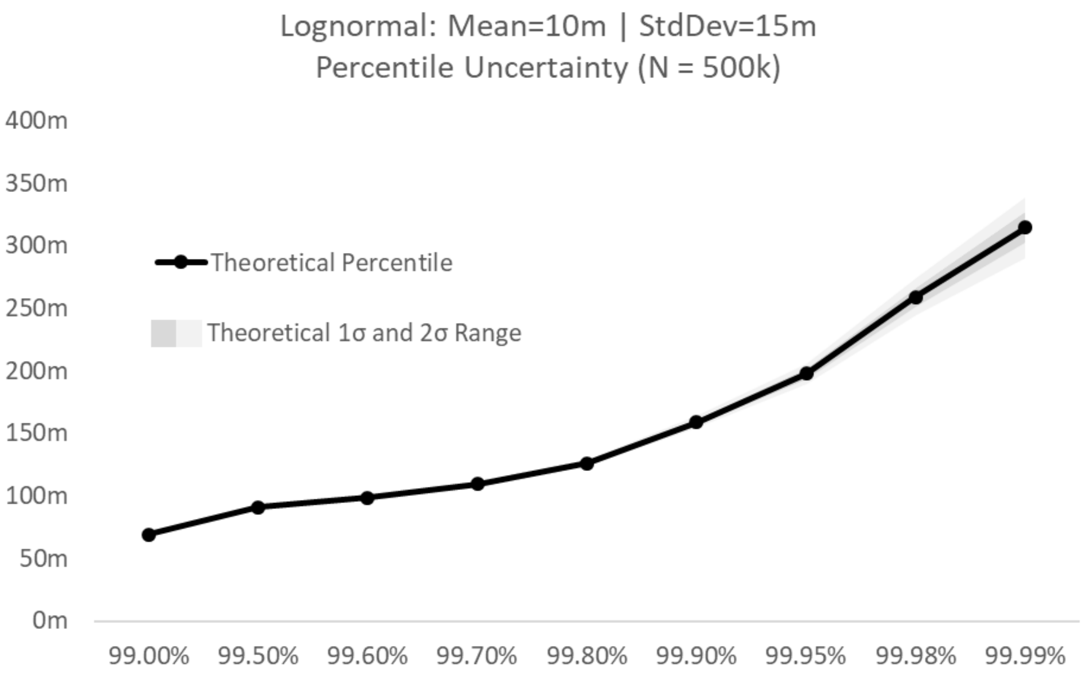



# Second-Order Uncertainty

## Volatility of Percentile Estimates: Known Distributions

- **There is a well-known asymptotic formula**: $\sigma \to \sqrt{\frac{p(1-p)}{samples}} \times F^{-1'}(p)$, where *F(x)* is the CDF of the underlying distribution, so $F^{-1}$ is the graph below and $F^{-1'}$ is just the gradient of the graph below at each percentile
- **For high percentiles**, our estimates are more reliable as the samples increase or as the gradient decreases
- **Lognormal example below**: Theory matches our experiment on Slide 13 above!



# Second-Order Uncertainty

## Volatility of Percentile Estimates: Unknown Distributions

- **But our simulations are crazy** – they don't have closed-form analytical distributions for $F^{-1'}(p)$!
- **So we use one simulation to estimate the gradient at each percentile**, focusing on its small neighborhood
- **Lognormal example below** to illustrate

# Second-Order Uncertainty

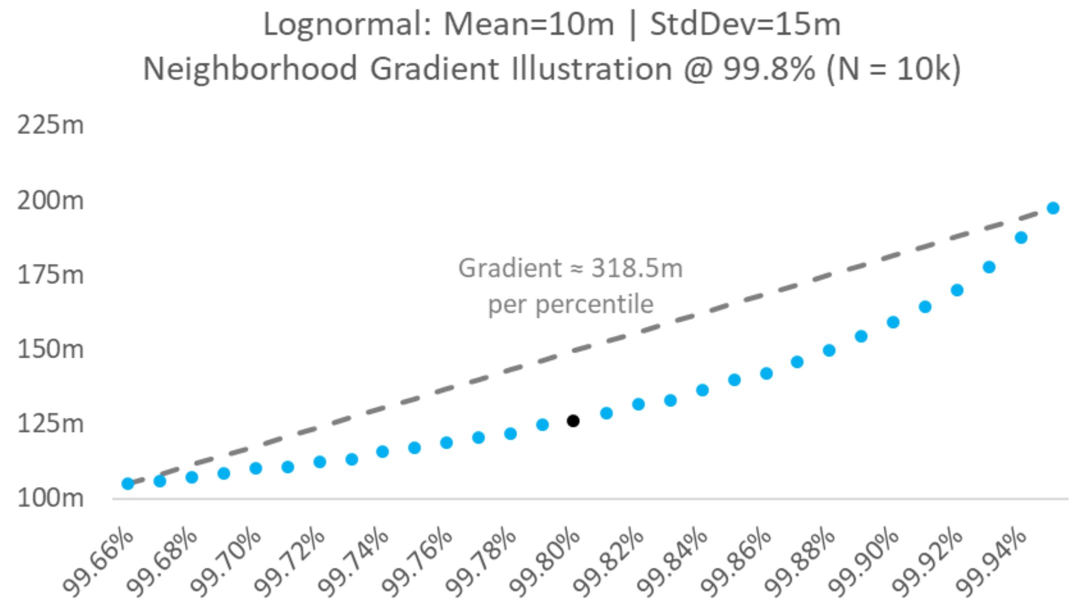
## Volatility of Percentile Estimates: Unknown Distributions

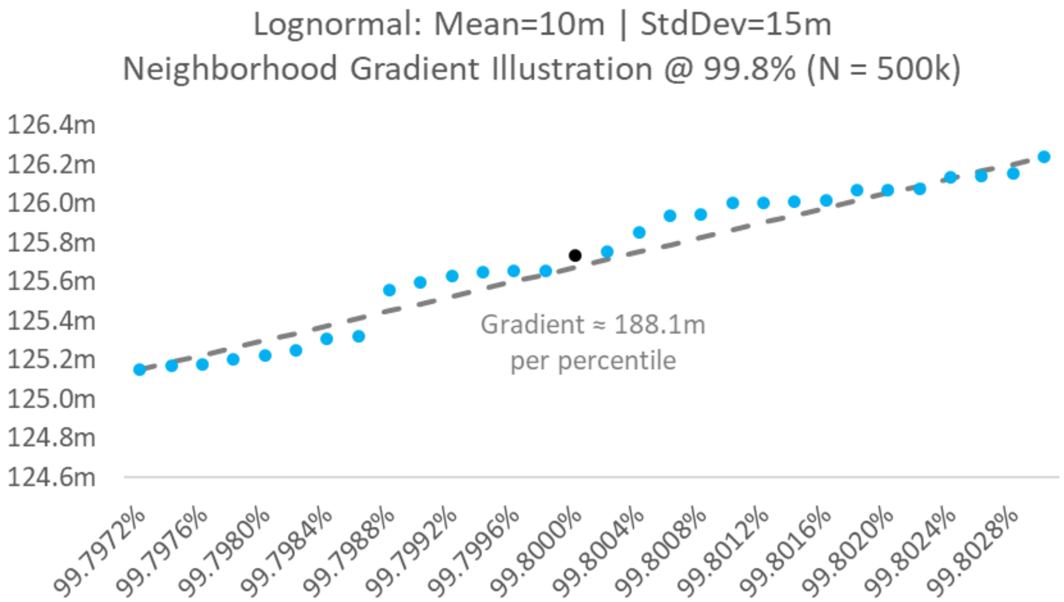
- **But our simulations are crazy** – they don't have closed-form analytical distributions for $F^{-1'}(p)$!
- **So we use one simulation to estimate the gradient at each percentile**, focusing on its small neighborhood
- **Lognormal example below** to illustrate

# Second-Order Uncertainty

## Volatility of Percentile Estimates: Unknown Distributions

- **But our simulations are crazy** – they don't have closed-form analytical distributions for ${F^{-1}}'(p)$!
- **So we use one simulation to estimate the gradient at each percentile**, focusing on its small neighborhood
- **Lognormal example below** to illustrate the estimation at the 99.8$^{th}$ percentile



# Second-Order Uncertainty

## Volatility of Percentile Estimates: Unknown Distributions

Instead of aiming for σ, we'll aim for the $CV = \frac{\sigma}{\mu}$ using the neighborhood of O observations around the percentile:

$$\widehat{CV} \to \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}} - \widehat{X_{low}}}{\hat{X}} = Multiplier \times Uncertainty\ \%\ in\ Obs\ Region$$

The CV

- As N grows, the Multiplier grows but Uncertainty % typically shrinks quicker due to the zoom-in effect
- Lognormal example below: Although the Multiplier increased with more samples, the Uncertainty % decreased by far more; so the 500k 99.8th percentile is more reliable than the 10k 99.8th, as expected!

# Second-Order Uncertainty

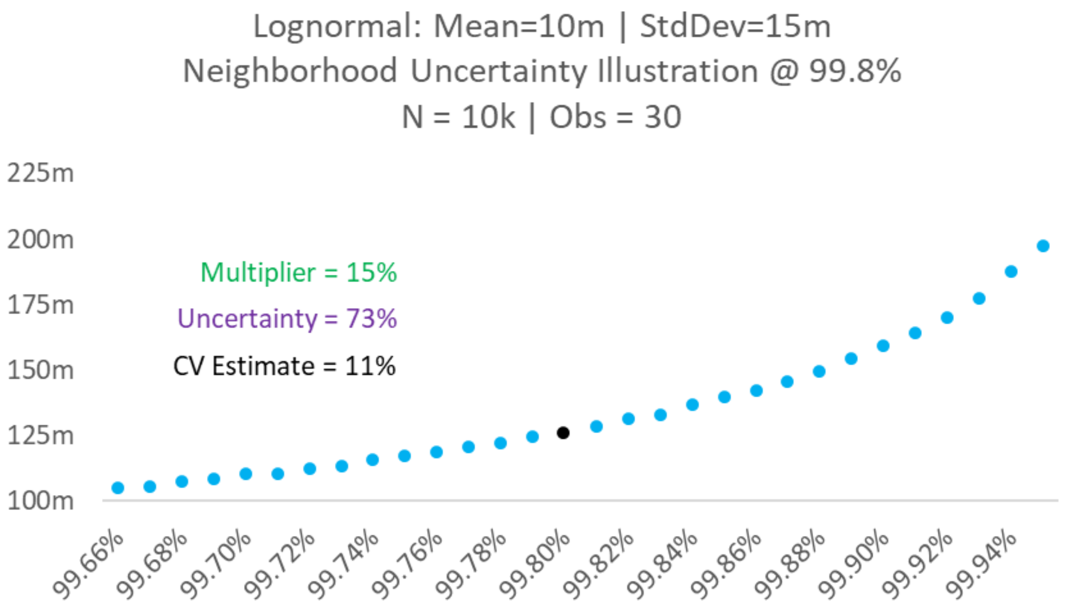
## Volatility of Percentile Estimates: Unknown Distributions

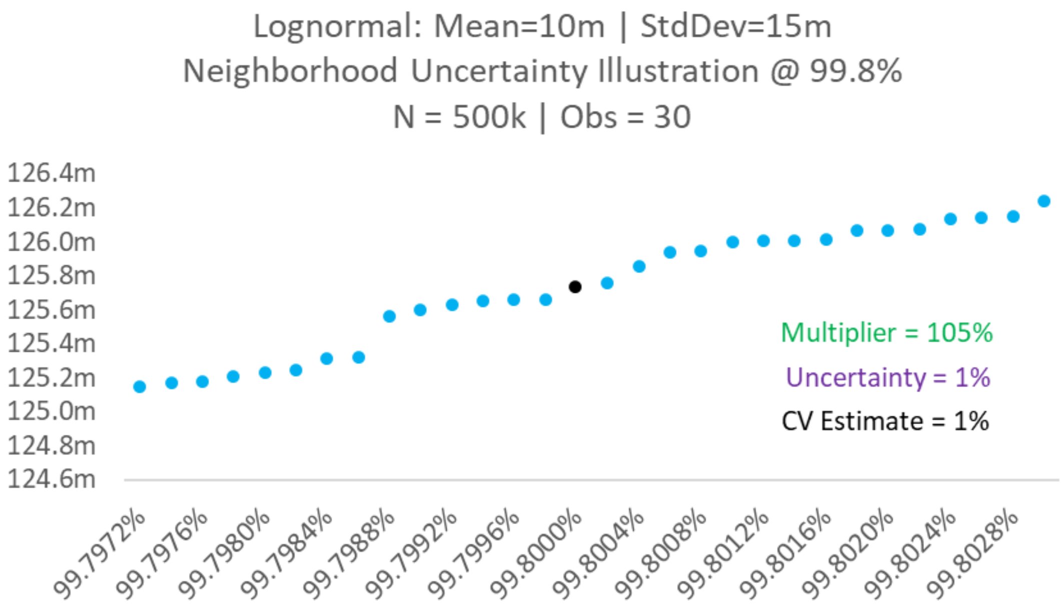
Instead of aiming for σ, we'll aim for the $CV = \frac{\sigma}{\mu}$ using the neighborhood of O observations around the percentile:

$$\widehat{CV} \rightarrow \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}} - \widehat{X_{low}}}{\hat{X}} = Multiplier \times Uncertainty\ \%\ in\ Obs\ Region$$

The CV

- As N grows, the Multiplier grows but Uncertainty % typically shrinks quicker due to the zoom-in effect
- Lognormal example below: Although the Multiplier increased with more samples, the Uncertainty % decreased by far more; so the 500k 99.8th percentile is more reliable than the 10k 99.8th, as expected!



# Second-Order Uncertainty

## A New Framework

Instead of aiming for σ, we'll aim for the $CV = \frac{\sigma}{\mu}$ using the neighborhood of O observations around the percentile:

$$\widehat{CV} \rightarrow \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}} - \widehat{X_{low}}}{\hat{X}} = Multiplier \times Uncertainty\ \%\ in\ Obs\ Region$$

So we can think of simulation uncertainty as having *three* levels:

1. **First-Order Uncertainty**: Any simulated metric (random variable) like gross losses, net losses, ceded to a contract, etc. has percentiles – i.e., **a CDF**
2. **Second-Order Uncertainty**: The simulated percentiles *themselves* have inherent uncertainty and will change if we re-simulate – i.e., *our attempt at* **Error Bars** using the imperfect information from our simulation
3. **Third-Order Uncertainty**: The *attempt at* Second-Order Uncertainty (we'll focus on the CV) will vary by simulation, depending on the Uncertainty % in the simulated neighborhood of observations around the percentile – i.e., **the Stretchiness/Blurriness of the Error Bars**

# Second-Order Uncertainty

## A New Framework

Instead of aiming for σ, we'll aim for the $CV = \frac{\sigma}{\mu}$ using the neighborhood of O observations around the percentile:

$$\widehat{CV} \rightarrow \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}} - \widehat{X_{low}}}{\hat{X}} = Multiplier \times Uncertainty\ \%\ in\ Obs\ Region$$

So we can think of simulation uncertainty as having *three* levels:

1. **First-Order Uncertainty**: Any simulated metric (random variable) like gross losses, net losses, ceded to a contract, etc. has percentiles – i.e., **a CDF**
2. **Second-Order Uncertainty**: The simulated percentiles *themselves* have inherent uncertainty and will change if we re-simulate – i.e., *our attempt at* **Error Bars** using the imperfect information from our simulation
3. **Third-Order Uncertainty**: The *attempt at* Second-Order Uncertainty (we'll focus on the CV) will vary by simulation, depending on the Uncertainty % in the simulated neighborhood of observations around the percentile – i.e., **the Stretchiness/Blurriness of the Error Bars**

# What’s the Deal with Third-Order Uncertainty?

# Third-Order Uncertainty

## Recap and New Direction

- When simulating a random variable we realized that the resulting percentiles (First-Order Uncertainty: CDF) *themselves* have uncertainty in the sense that they will change on re-simulation – i.e., they have Second-Order Uncertainty (Error Bars)
- Luckily, though, we have a nice formula to calculate the CV of any simulated percentile:

$$\widehat{CV} \to \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}} - \widehat{X_{low}}}{\hat{X}} = Multiplier \times Uncertainty\ \%\ in\ Obs\ Region$$

- But *even that* is uncertain, as it will change on re-simulation depending on the Uncertainty % in the neighborhood. So the natural question to ask is… How reliable is our $\widehat{CV}$? That is, how much should we expect this $\widehat{CV}$ to change each time we re-simulate? Or 'How stretchy/blurry are our error bars?'

We will find the **CV of $\widehat{CV}$ at any given percentile**! This will be our measure of Third-Order Uncertainty, and the first place to look is in some simulated results on well-understood distributions

# Third-Order Uncertainty

## Recap and New Direction

- When simulating a random variable we realized that the resulting percentiles (First-Order Uncertainty: CDF) *themselves* have uncertainty in the sense that they will change on re-simulation – i.e., they have Second-Order Uncertainty (Error Bars)
- Luckily, though, we have a nice formula to calculate the CV of any simulated percentile:

$$\widehat{CV} \to \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}} - \widehat{X_{low}}}{\hat{X}} = Multiplier \times Uncertainty\ \%\ in\ Obs\ Region$$

- But *even that* is uncertain, as it will change on re-simulation depending on the Uncertainty % in the neighborhood. So the natural question to ask is… How reliable is our $\widehat{CV}$? That is, how much should we expect this $\widehat{CV}$ to change each time we re-simulate? Or 'How stretchy/blurry are our error bars?'

We will find the **CV of $\widehat{CV}$ at any given percentile**! This will be our measure of Third-Order Uncertainty, and the first place to look is in some simulated results on well-understood distributions

# Third-Order Uncertainty

## Recap and New Direction

- When simulating a random variable we realized that the resulting percentiles (First-Order Uncertainty: CDF) *themselves* have uncertainty in the sense that they will change on re-simulation – i.e., they have Second-Order Uncertainty (Error Bars)
- Luckily, though, we have a nice formula to calculate the CV of any simulated percentile:

$$\widehat{CV} \to \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}} - \widehat{X_{low}}}{\hat{X}} = Multiplier \times Uncertainty\ \%\ in\ Obs\ Region$$

- But *even that* is uncertain, as it will change on re-simulation depending on the Uncertainty % in the neighborhood. So the natural question to ask is… How reliable is our $\widehat{CV}$? That is, how much should we expect this $\widehat{CV}$ to change each time we re-simulate? Or 'How stretchy/blurry are our error bars?'

We will find the **CV of $\widehat{CV}$ at any given percentile**! This will be our measure of Third-Order Uncertainty, and the first place to look is in some simulated results on well-understood distributions

# Third-Order Uncertainty

## Recap and New Direction

- When simulating a random variable we realized that the resulting percentiles (First-Order Uncertainty: CDF) *themselves* have uncertainty in the sense that they will change on re-simulation – i.e., they have Second-Order Uncertainty (Error Bars)
- Luckily, though, we have a nice formula to calculate the CV of any simulated percentile:

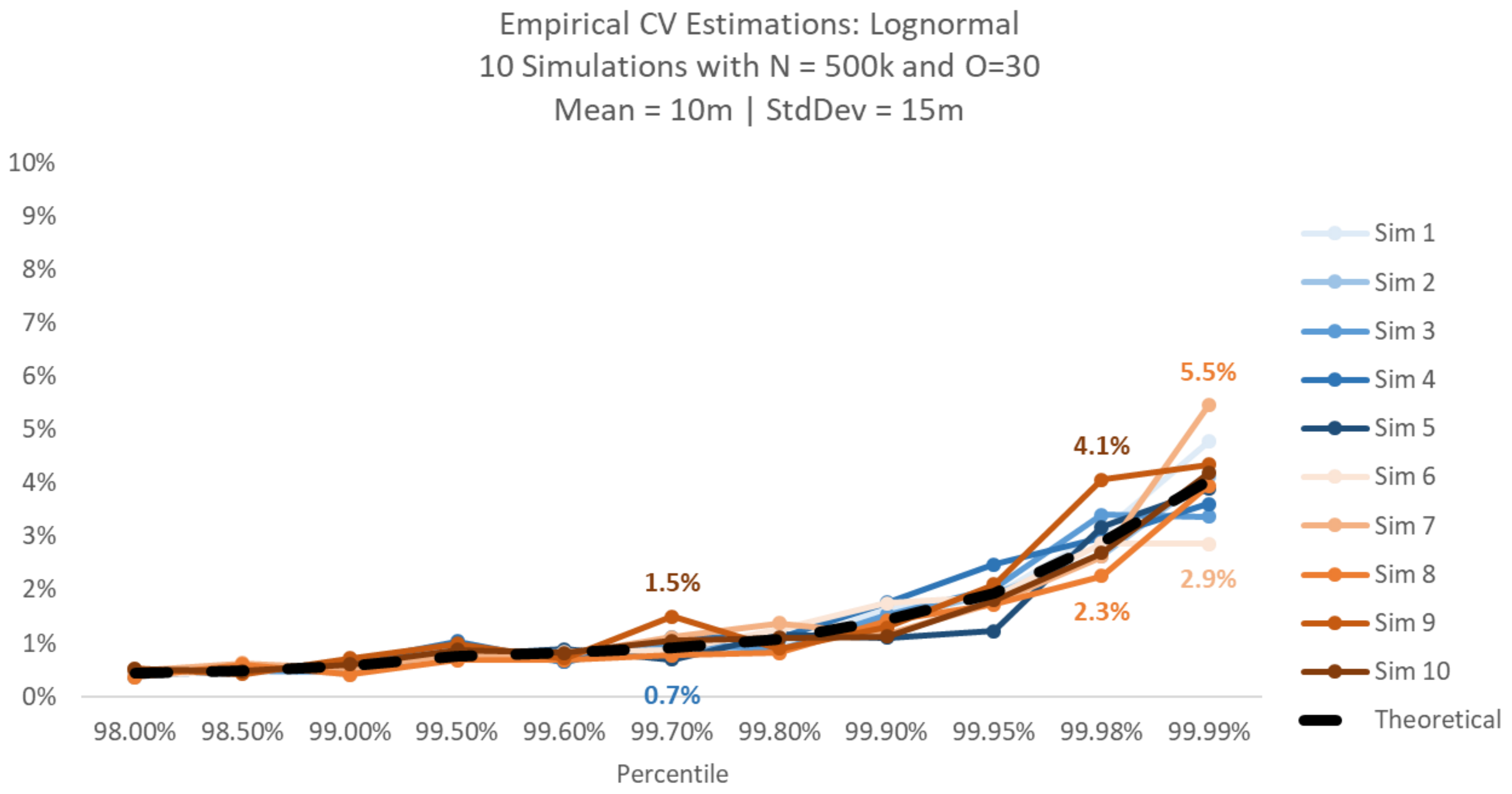
$$\widehat{CV} \rightarrow \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}} - \widehat{X_{low}}}{\hat{X}} = Multiplier \times Uncertainty\ \%\ in\ Obs\ Region$$

- But *even that* is uncertain, as it will change on re-simulation depending on the Uncertainty % in the neighborhood. So the natural question to ask is… How reliable is our $\widehat{CV}$? That is, how much should we expect this $\widehat{CV}$ to change each time we re-simulate? Or 'How stretchy/blurry are our error bars?'

We will find the **CV of $\widehat{CV}$ at any given percentile**! This will be our measure of Third-Order Uncertainty, and the first place to look is in some simulated results on well-understood distributions

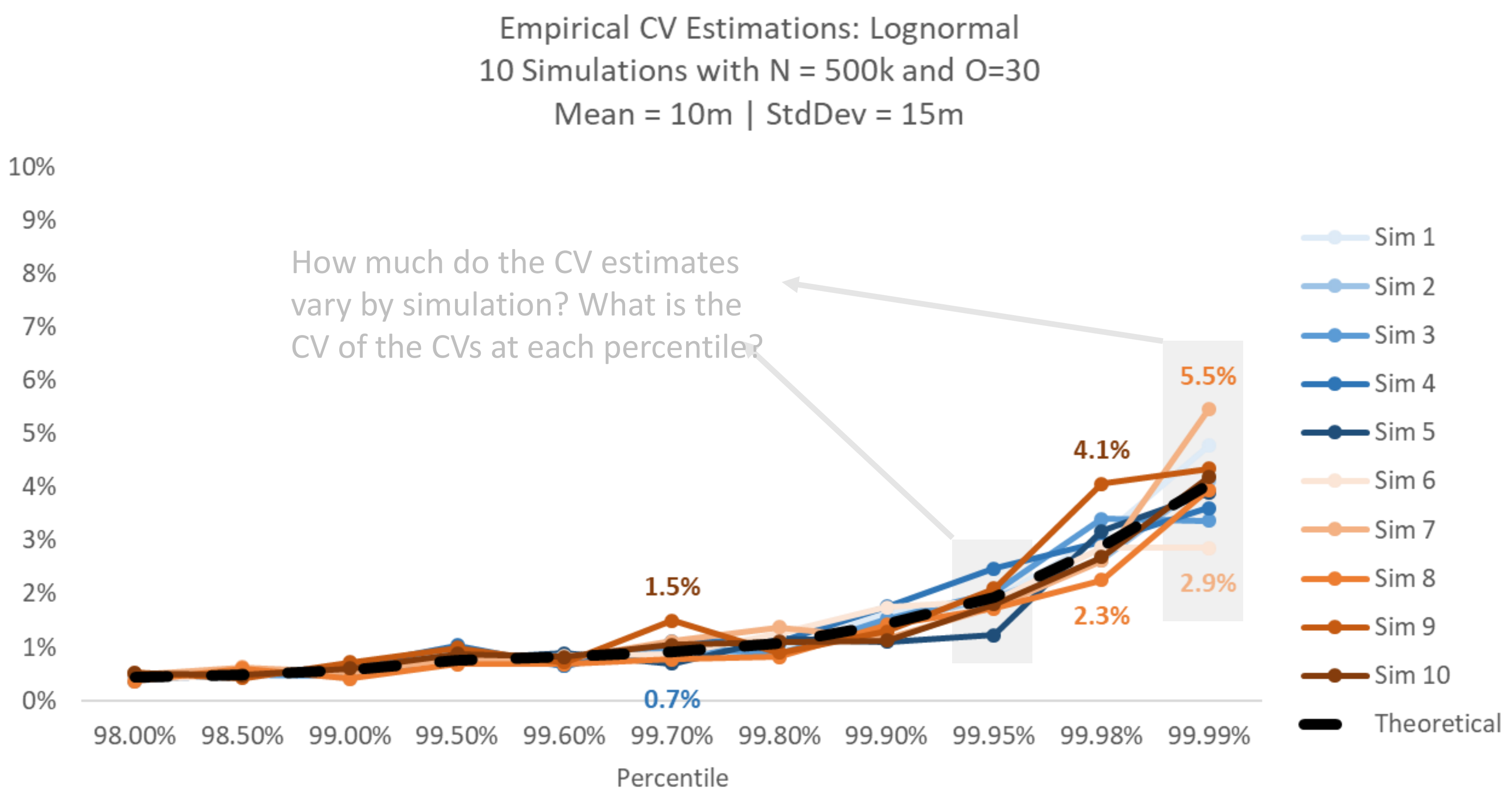
# Third-Order Uncertainty

## Empirical Behavior: 10 Sims of a Lognormal Percentile's CVs



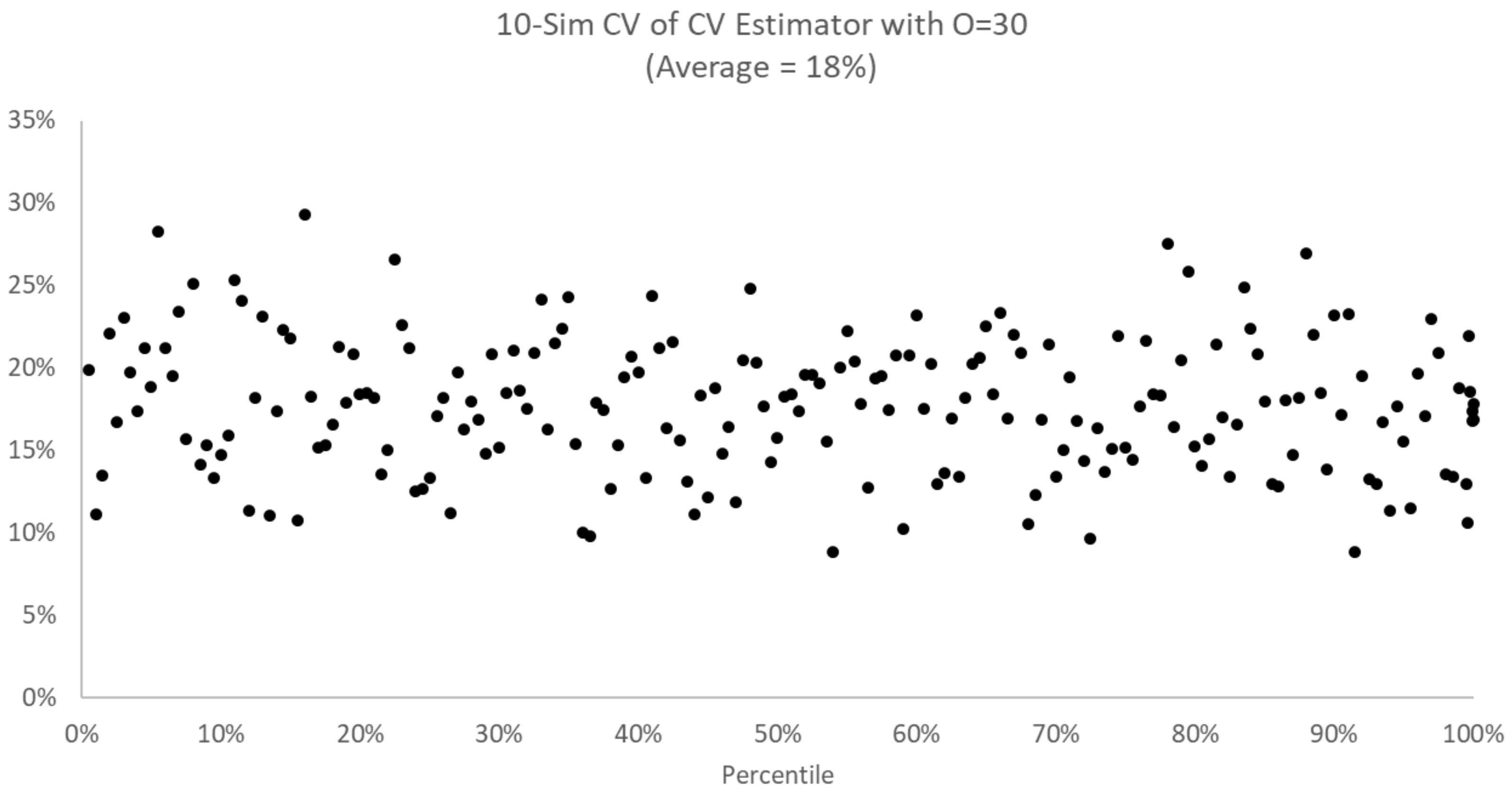
# Third-Order Uncertainty

## Empirical Behavior: 10 Sims of a Lognormal Percentile's CVs



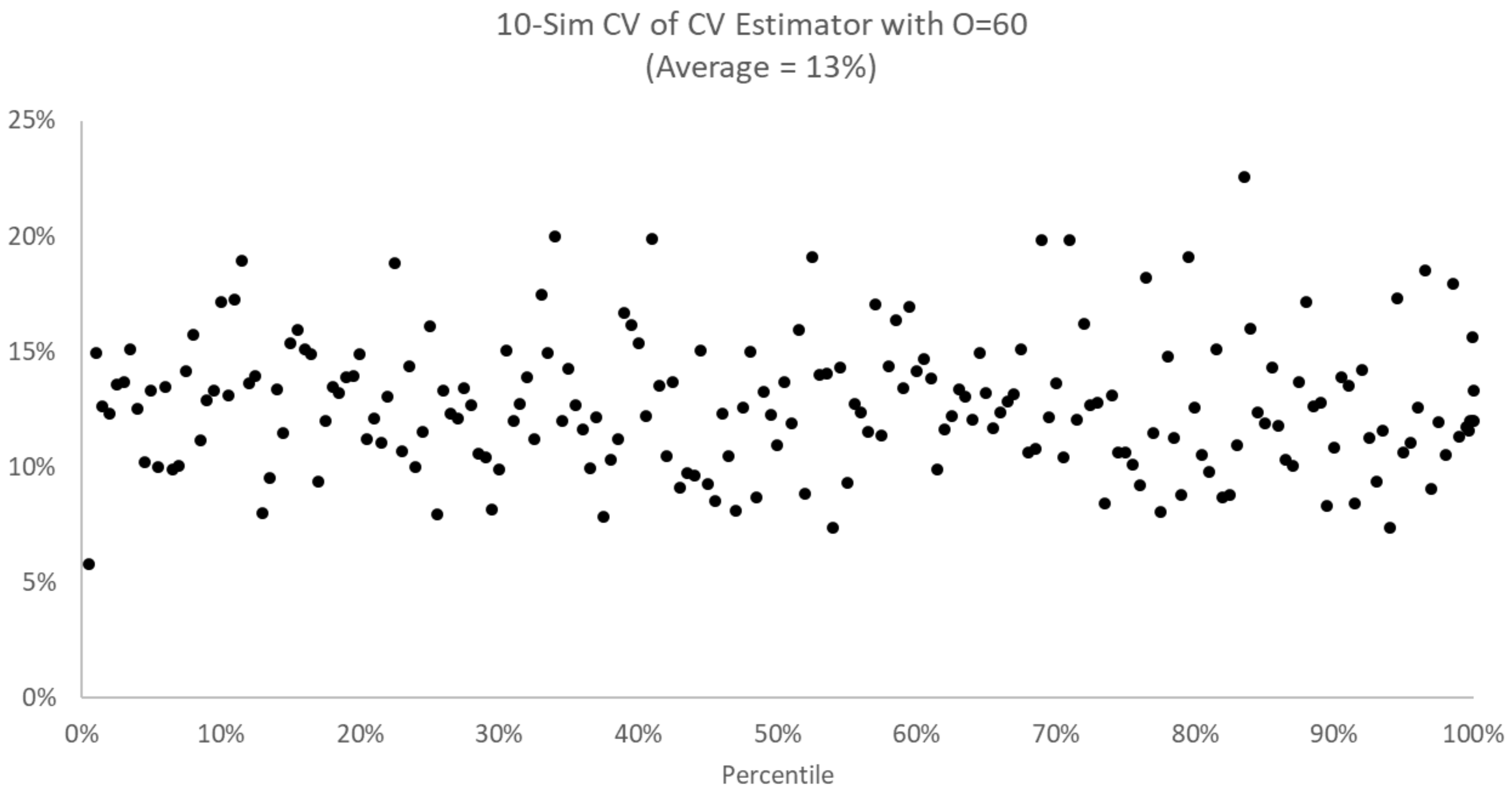
# Third-Order Uncertainty

## Empirical Behavior: 10 Sims of a Lognormal Percentile's CVs



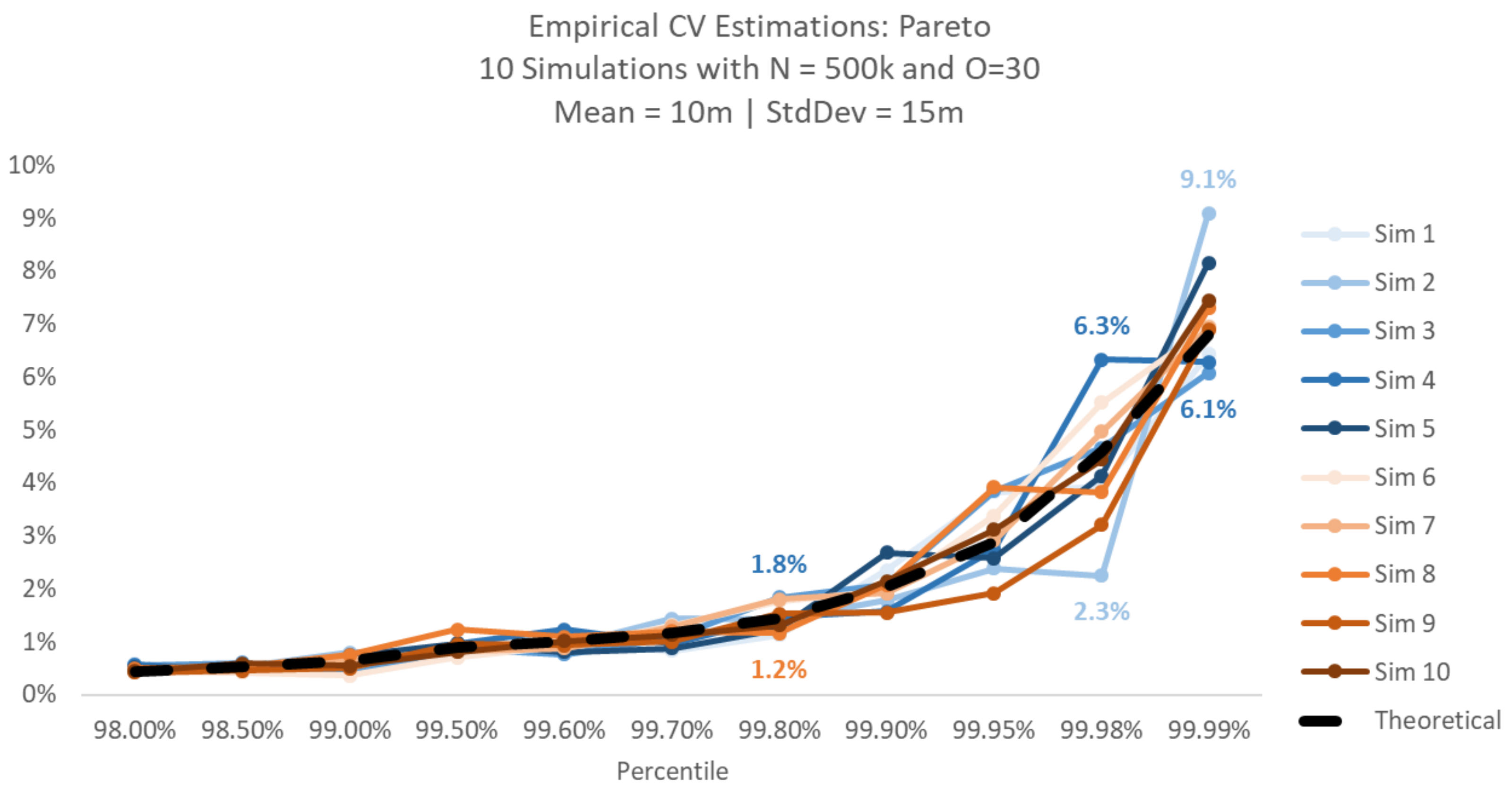
# Third-Order Uncertainty

## Empirical Behavior: 10 Sims of a Lognormal Percentile’s CVs



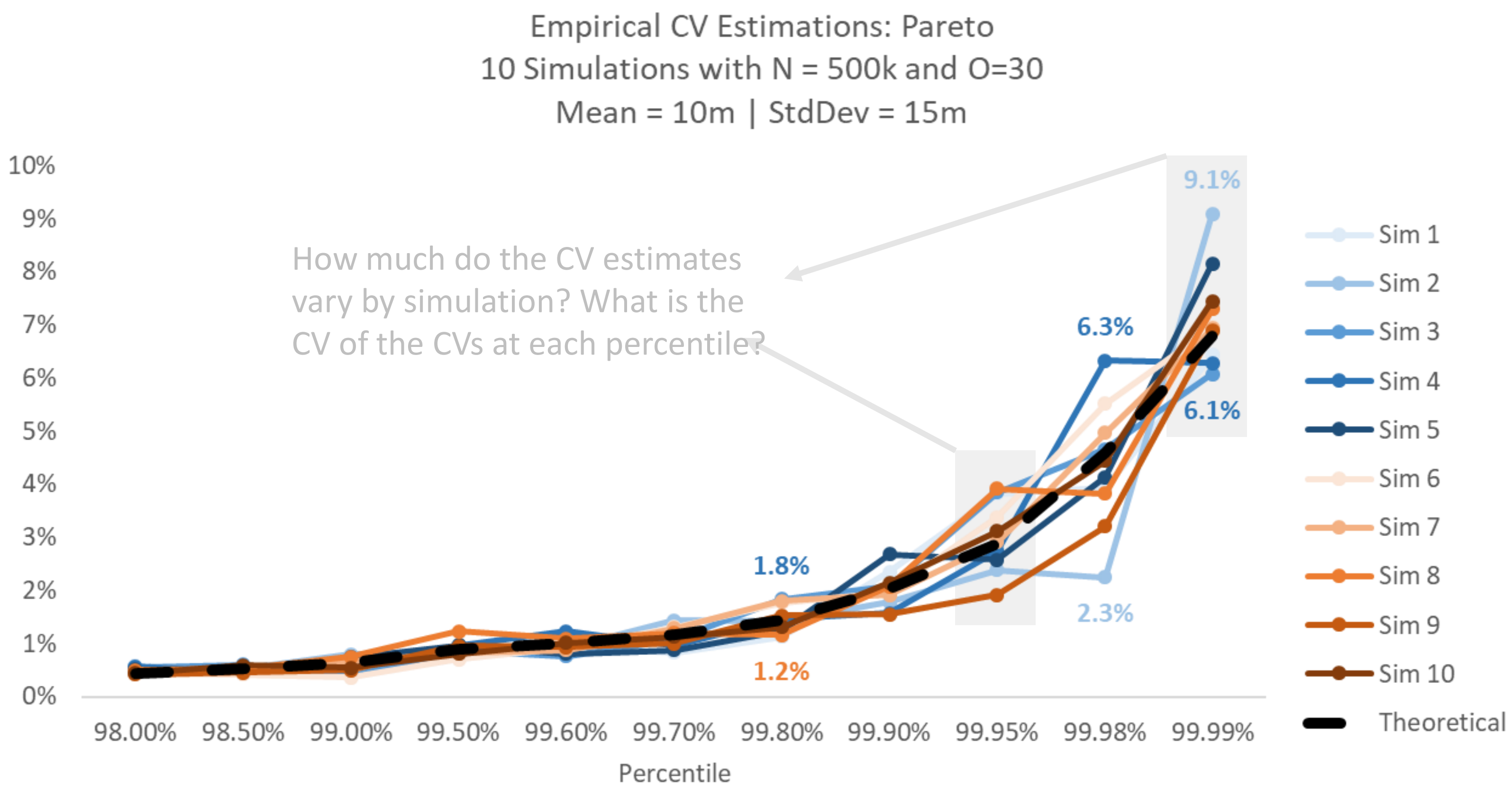
# Third-Order Uncertainty

## Empirical Behavior: 10 Sims of a Pareto Percentile's CVs



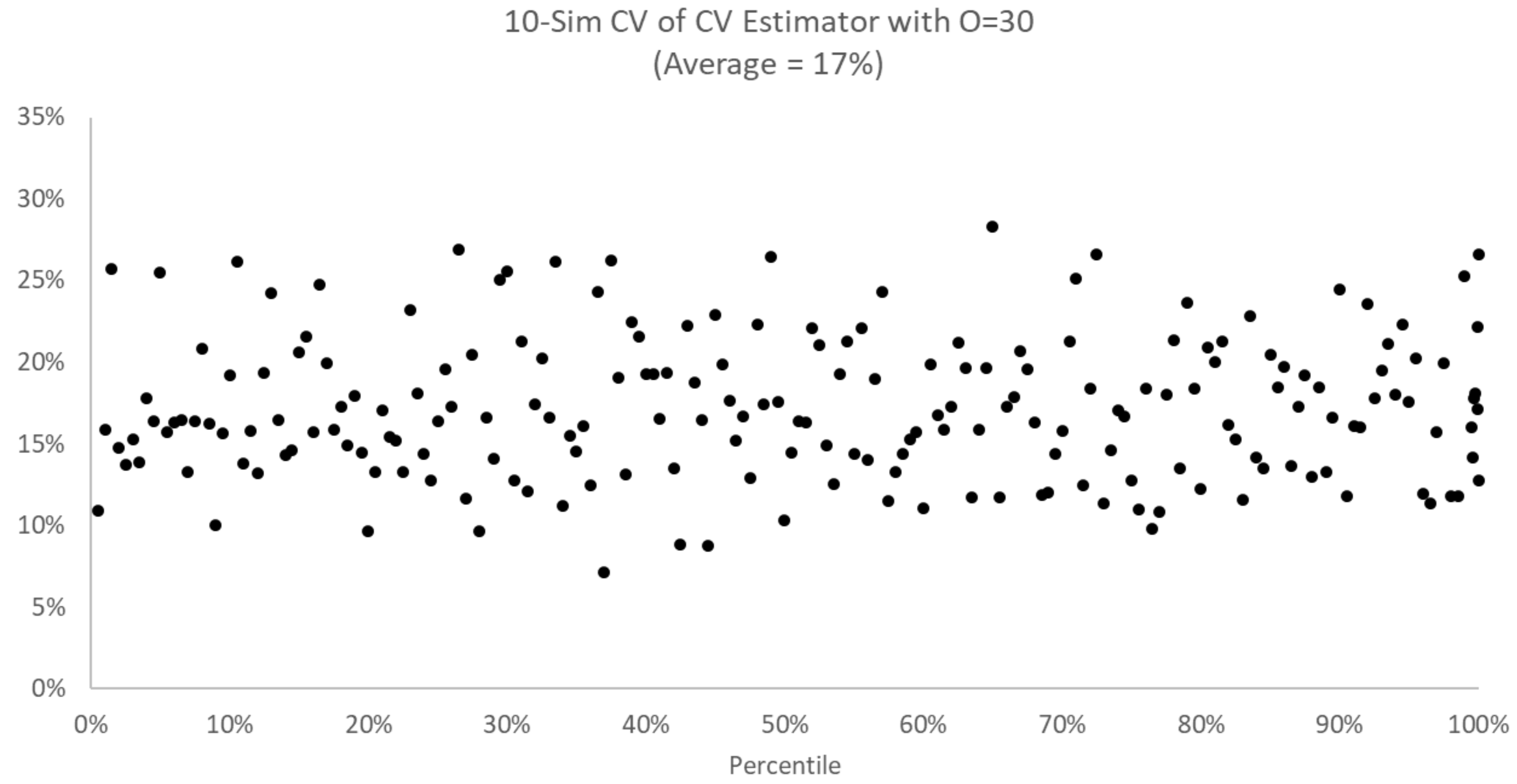
# Third-Order Uncertainty

## Empirical Behavior: 10 Sims of a Pareto Percentile's CVs



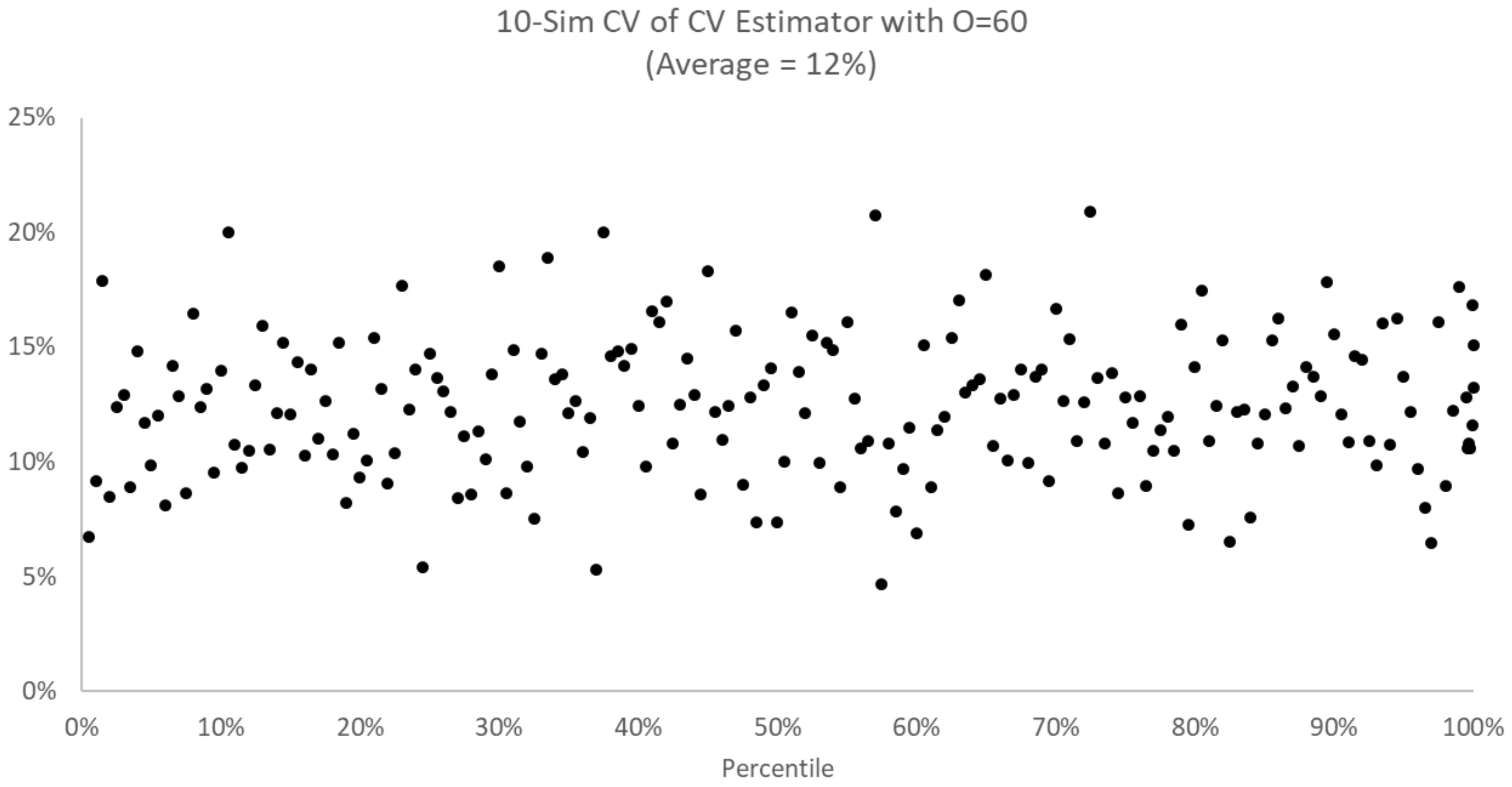
# Third-Order Uncertainty

## Empirical Behavior: 10 Sims of a Pareto Percentile’s CVs



# Third-Order Uncertainty

## Empirical Behavior: 10 Sims of a Pareto Percentile's CVs

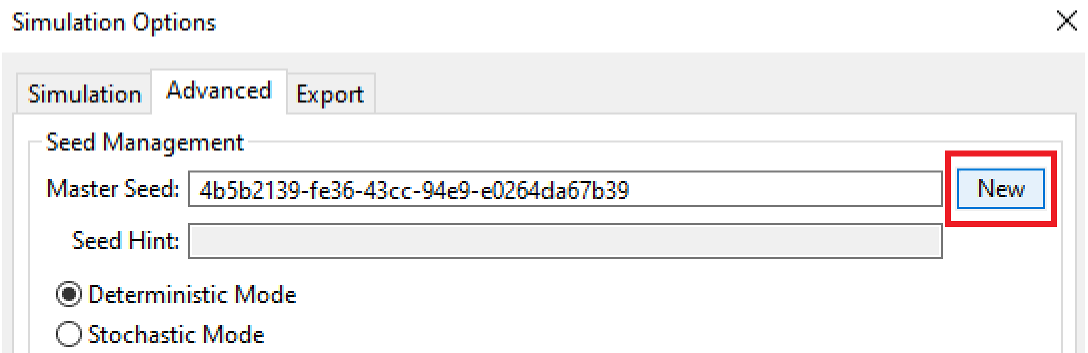


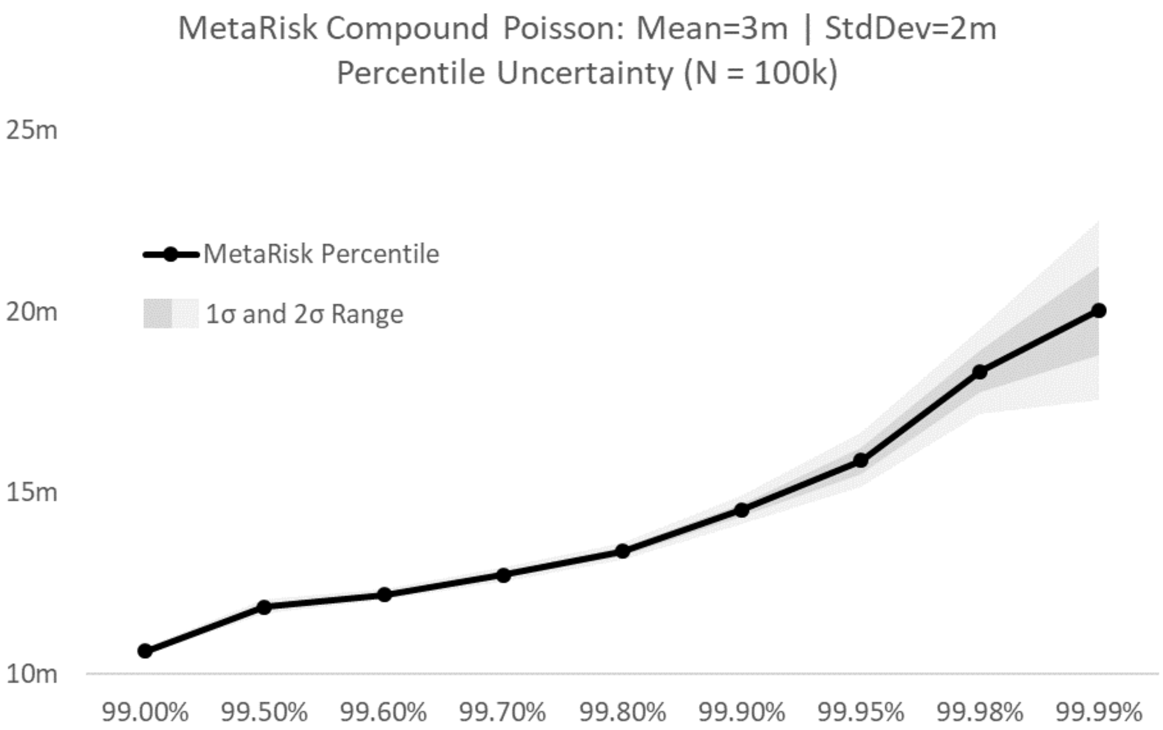
# Third-Order Uncertainty
## Key Results

1. **The $\widehat{CV}$ *of percentile* $p$** $\rightarrow \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}} - \widehat{X_{low}}}{\hat{X}} =$ $Multiplier \times Uncertainty\ \%\ in\ Obs\ Region$

2. **The percentiles' $\widehat{CV}$s *themselves* vary** across simulations due to the randomness in $Uncertainty\ \%$ : $\mathbf{CV}(\widehat{CV}) \approx \frac{1}{\sqrt{O}}$ *regardless of the percentile or the distribution* (Proof in Appendix), so $\widehat{CV}$ is only reliable with many realizations and a big Observation Region on a well-behaved distribution

As an example, the error bands on the right might stretch up/down around 20% if I run a new simulation with O = 30!



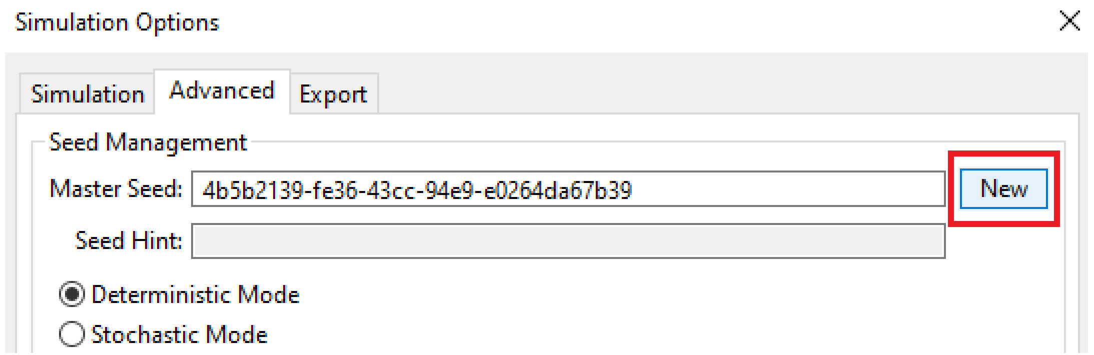
# Third-Order Uncertainty

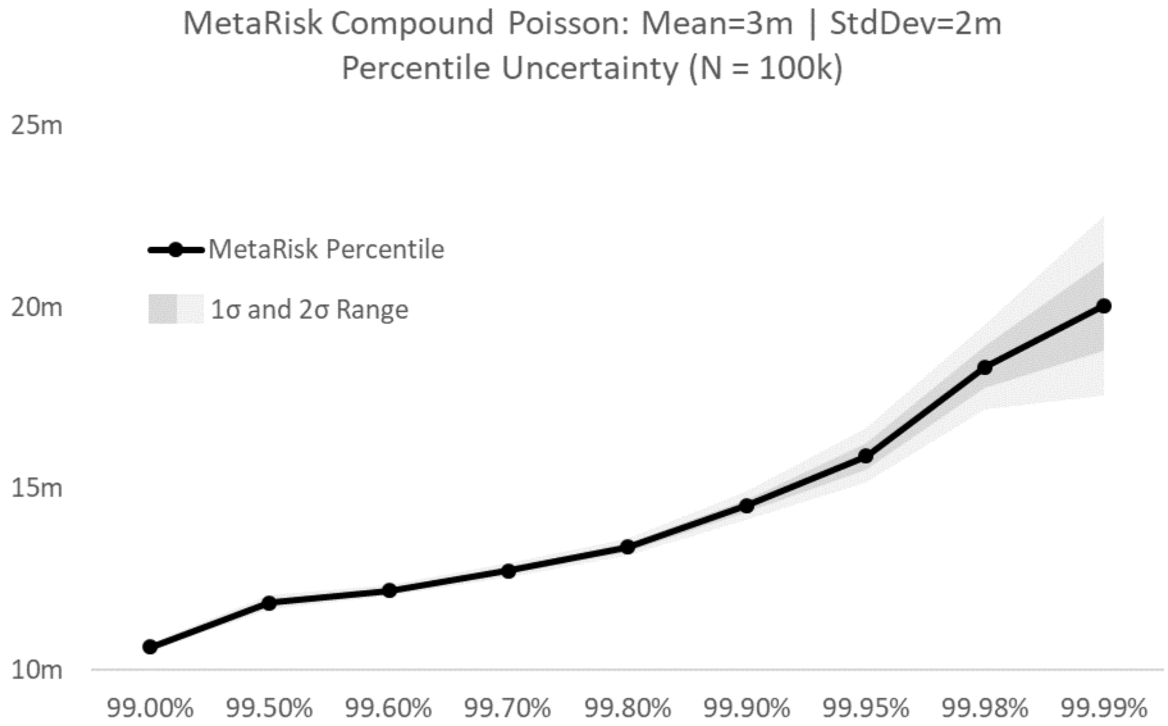
## Key Results

1. **The** $\widehat{CV}$ ***of percentile*** $p \rightarrow \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}}-\widehat{X_{low}}}{\hat{X}} = \mathit{Multiplier} \times \mathit{Uncertainty\ \%\ in\ Obs\ Region}$

2. **The percentiles'** $\widehat{CV}$**s *themselves* vary** across simulations due to the randomness in $\mathit{Uncertainty\ \%}$ : $\mathbf{CV}(\widehat{CV}) \approx \frac{1}{\sqrt{O}}$ *regardless of the percentile or the distribution* (Proof in Appendix), so $\widehat{CV}$ is only reliable with many realizations and a big Observation Region on a well-behaved distribution

As an example, the error bands on the right might stretch up/down around 20% if I run a new simulation with O = 30!



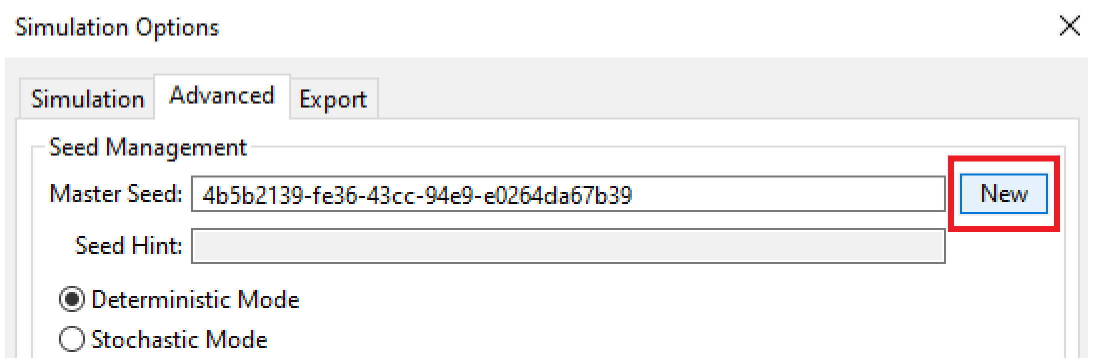
# Third-Order Uncertainty

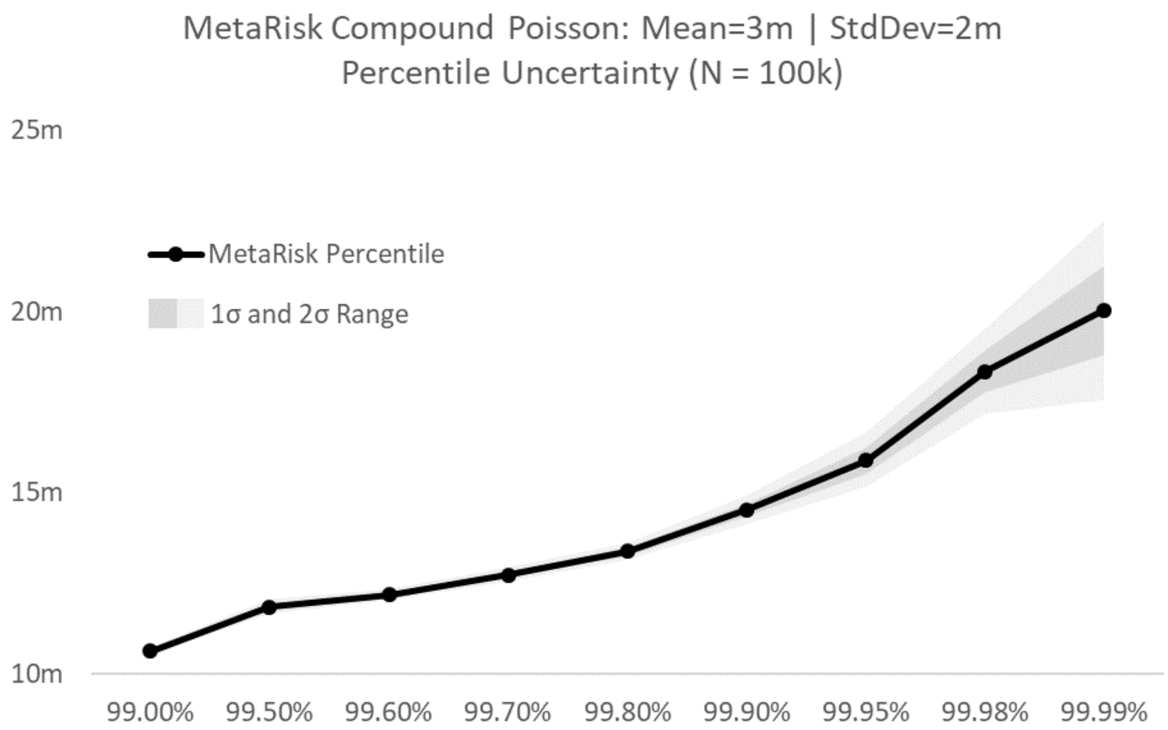
## Key Results

1. **The $\widehat{CV}$ *of percentile p*** $\rightarrow \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}} - \widehat{X_{low}}}{\hat{X}} =$ $Multiplier \times Uncertainty\ \%\ in\ Obs\ Region$

2. **The percentiles' $\widehat{CV}$s *themselves* vary** across simulations due to the randomness in $Uncertainty\ \%$ : $\mathbf{CV}(\widehat{CV}) \approx \frac{1}{\sqrt{O}}$ *regardless of the percentile or the distribution* (Proof in Appendix), so $\widehat{CV}$ is only reliable with many realizations and a big Observation Region on a well-behaved distribution

As an example, the error bands on the right might stretch up/down around 20% if I run a new simulation with O = 30!



# 5. Cat, Quota Share and PPR Case Studies: How Reliable are Our Percentiles?

# How Reliable are Our Percentiles?

## Two Questions in Practice

All results so far are asymptotic for well-behaved distributions – it's unclear whether the results hold at all for our simulations in the sub-10m realization regime.

1. **If the CV is small enough then it doesn't really matter how reliable it is!** That is, if the asymptotic Second-Order Uncertainty is ≈0.5%, then I'm happy with my simulated results and don't really care about my empirical or theoretical Second or Third-Order Uncertainty estimate.

2. **If the CV isn't small, then is it at least reliably close to $\widehat{CV}$?** In particular, if we re-simulate with a new seed:
    i. Are the percentiles actually within around *2σ* of the initially estimated CV?
    ii. Are the new estimated CVs wildly different to the initial ones – i.e., way more than $\frac{1}{\sqrt{Obs\ in\ Hood}}$ off?

# How Reliable are Our Percentiles?

## Two Questions in Practice

All results so far are asymptotic for well-behaved distributions – it's unclear whether the results hold at all for our simulations in the sub-10m realization regime.

1. **If the CV is small enough then it doesn't really matter how reliable it is!** That is, if the asymptotic Second-Order Uncertainty is ≈0.5%, then I'm happy with my simulated results and don't really care about my empirical or theoretical Second or Third-Order Uncertainty estimate.

2. **If the CV isn't small, then is it at least reliably close to $\widehat{CV}$?** In particular, if we re-simulate with a new seed:
    i. Are the percentiles actually within around $2\sigma$ of the initially estimated CV?
    ii. Are the new estimated CVs wildly different to the initial ones – i.e., way more than $\frac{1}{\sqrt{Obs\ in\ Hood}}$ off?

# How Reliable are Our Percentiles?

## Two Questions in Practice

All results so far are asymptotic for well-behaved distributions – it's unclear whether the results hold at all for our simulations in the sub-10m realization regime.

1. **If the CV is small enough then it doesn't really matter how reliable it is!** That is, if the asymptotic Second-Order Uncertainty is ≈0.5%, then I'm happy with my simulated results and don't really care about my empirical or theoretical Second or Third-Order Uncertainty estimate.

2. **If the CV isn't small, then is it at least reliably close to $\widehat{CV}$?** In particular, if we re-simulate with a new seed:
    i. Are the percentiles actually within around *2σ* of the initially estimated CV?
    ii. Are the new estimated CVs wildly different to the initial ones – i.e., way more than $\frac{1}{\sqrt{Obs\ in\ Hood}}$ off?

# How Reliable are Our Percentiles?

## Two Questions in Practice

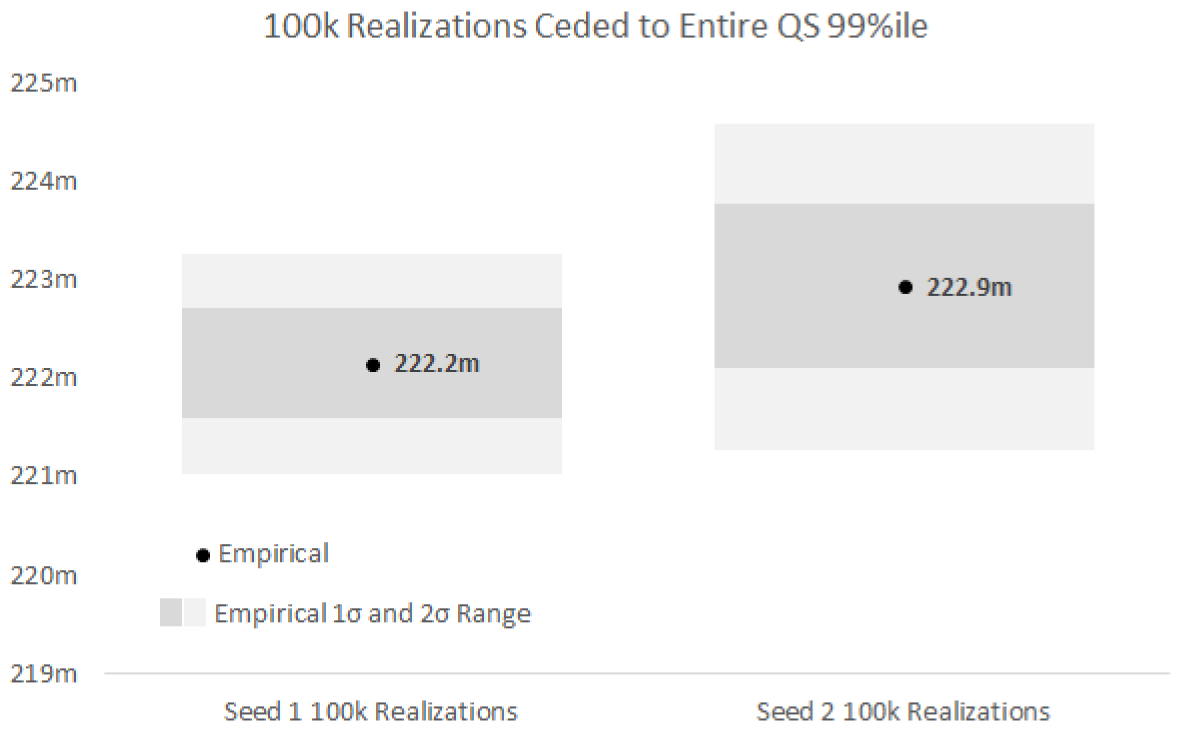
All results so far are asymptotic for well-behaved distributions – it's unclear whether the results hold at all for our simulations in the sub-10m realization regime.

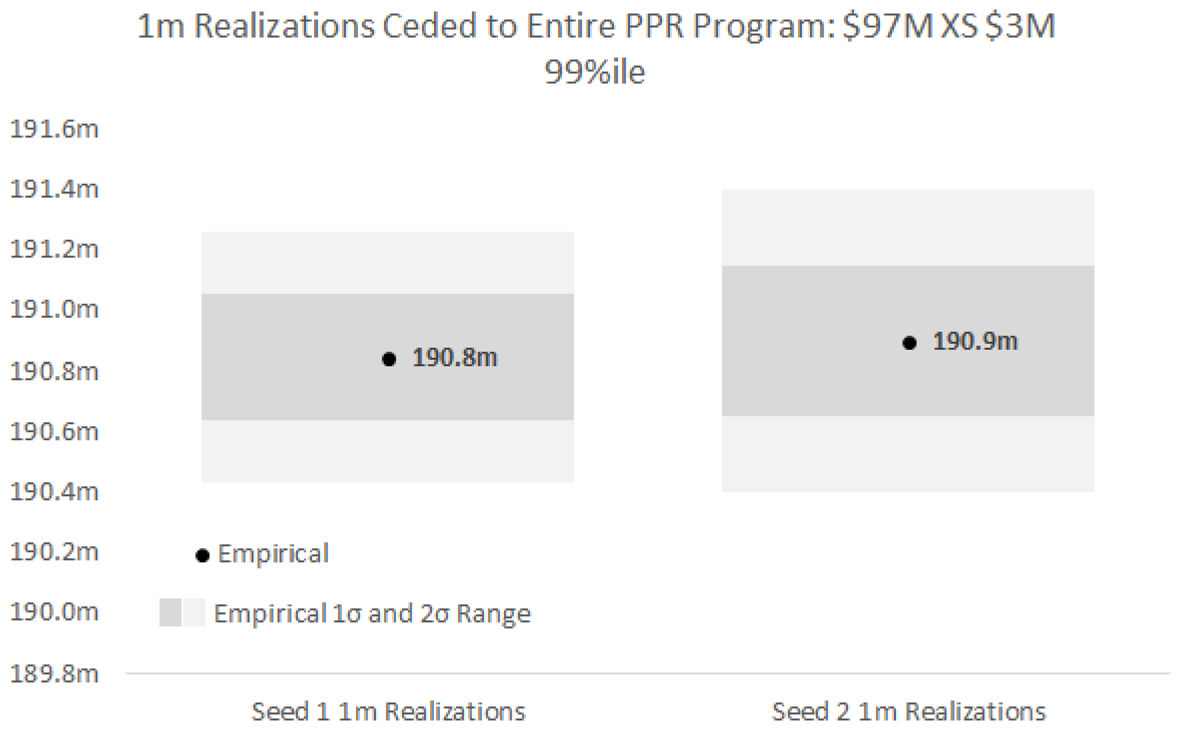
1. **If the CV is small enough then it doesn't really matter how reliable it is!** That is, if the asymptotic Second-Order Uncertainty is ≈0.5%, then I'm happy with my simulated results and don't really care about my empirical or theoretical Second or Third-Order Uncertainty estimate.

2. **If the CV isn't small, then is it at least reliably close to $\widehat{CV}$?** In particular, if we re-simulate with a new seed:
    i. Are the percentiles actually within around *2σ* of the initially estimated CV?
    ii. Are the new estimated CVs wildly different to the initial ones – i.e., way more than $\frac{1}{\sqrt{Obs\ in\ Hood}}$ off?

# How Reliable are Our Percentiles?

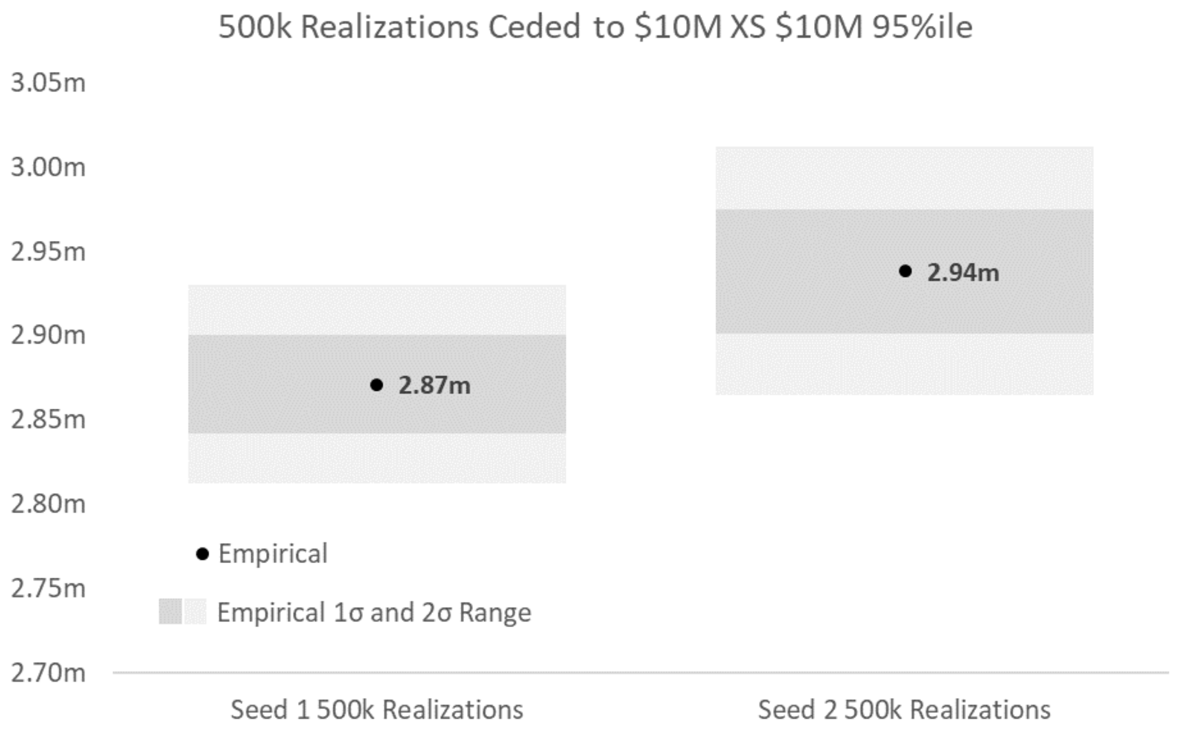
## Quota Share and Prop Per Risk Using RMS Seem Good!

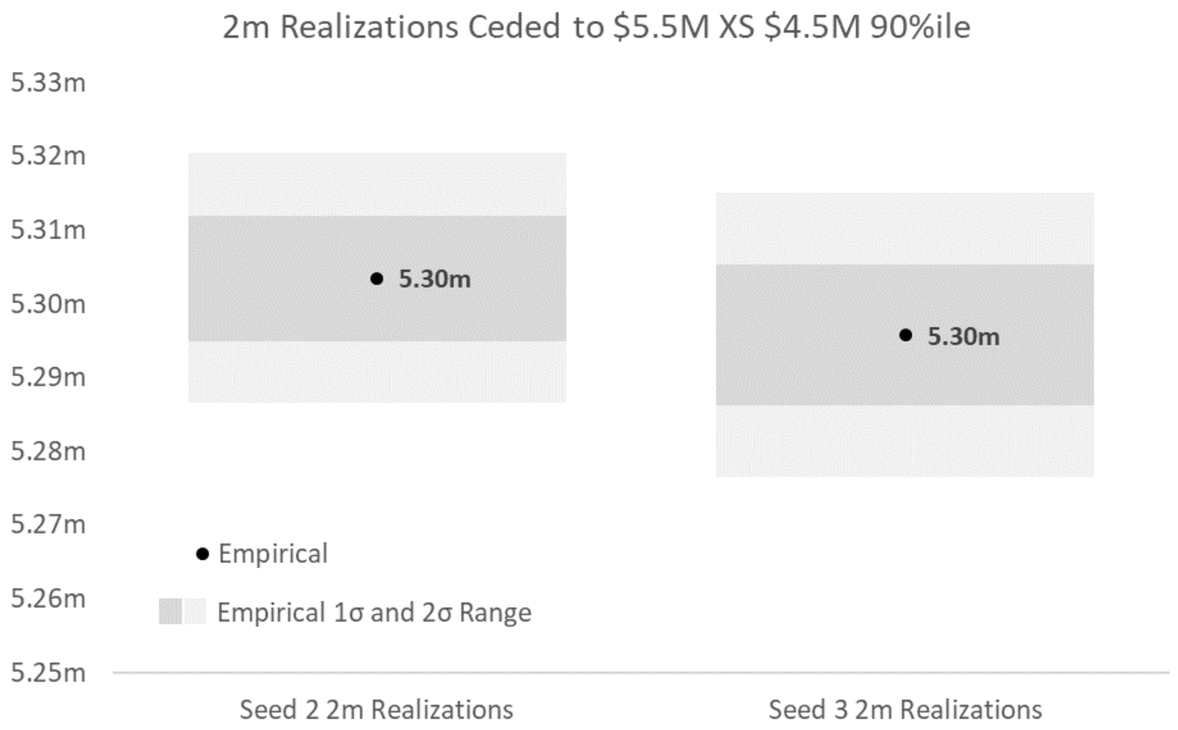


# How Reliable are Our Percentiles?

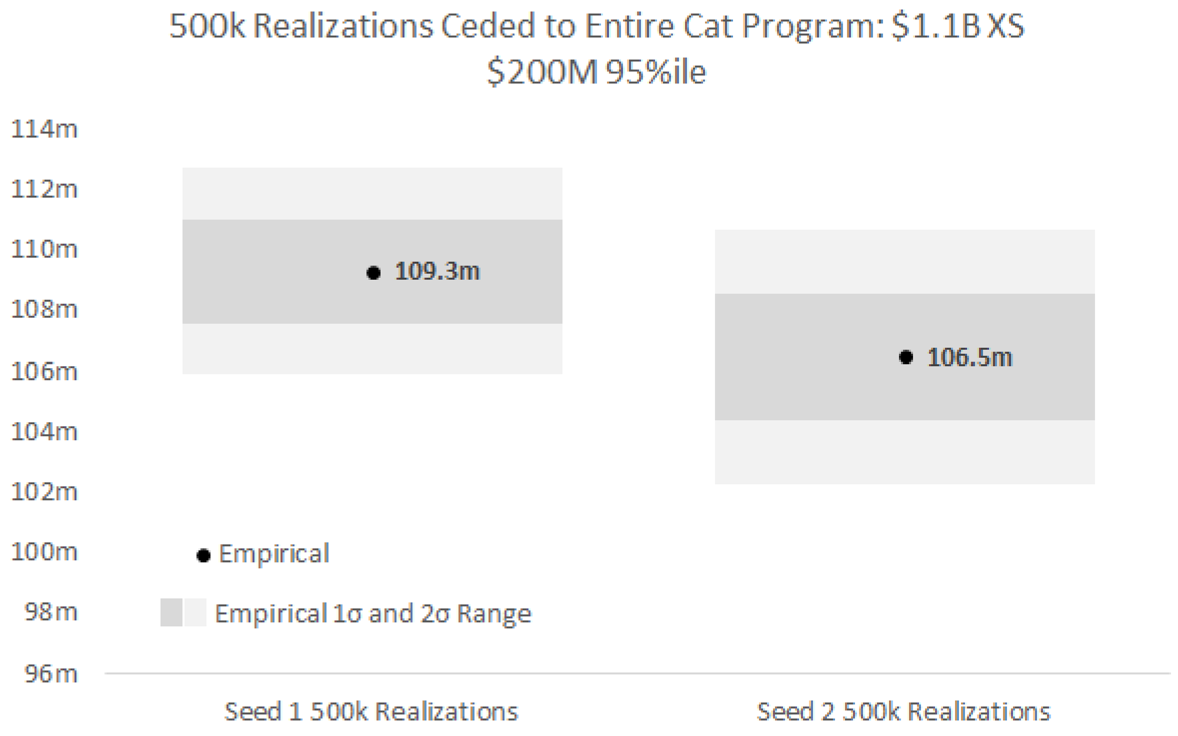
## Workers' Comp XOL incl. Cat (RMS) Might be Good!

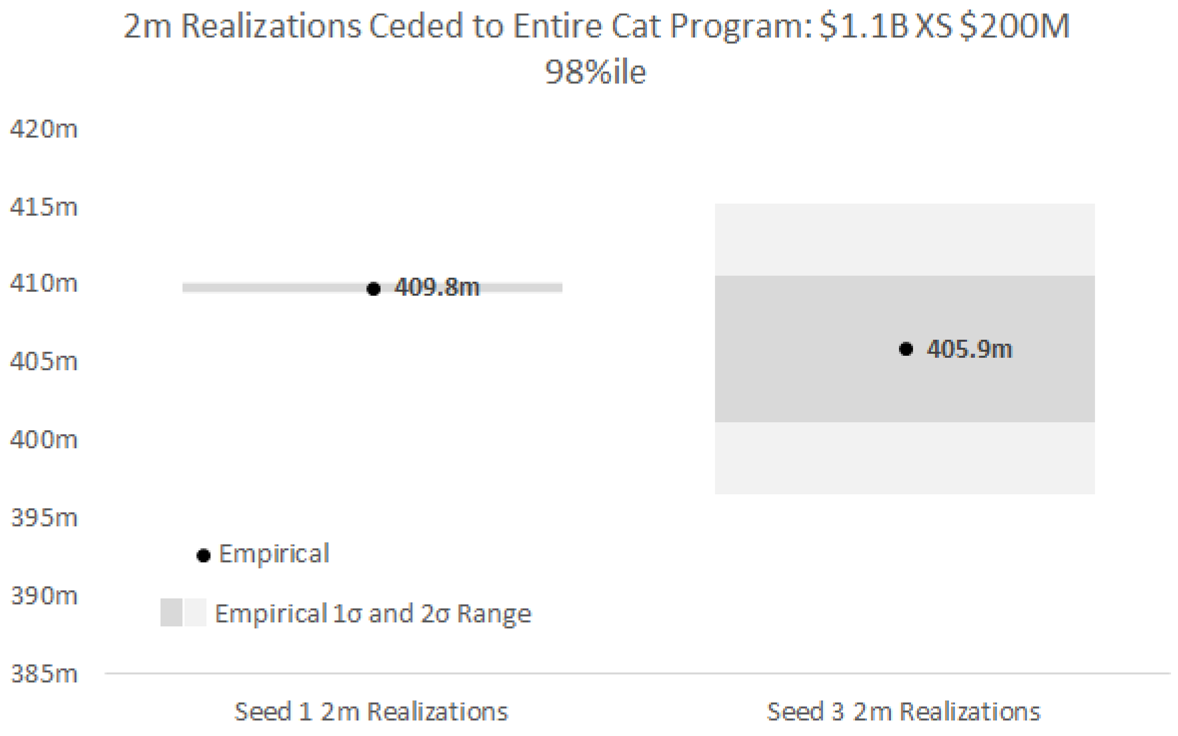


# How Reliable are Our Percentiles?

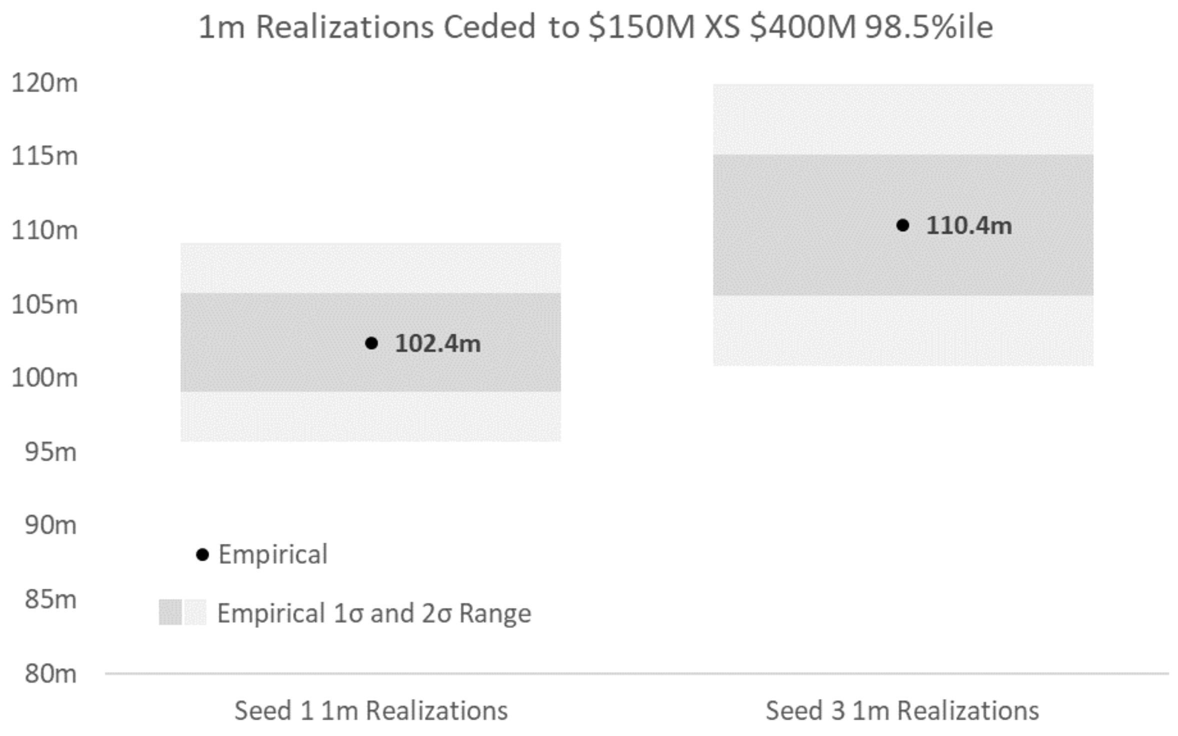
## High Excess Cat Using RMS is Problematic: Client 1

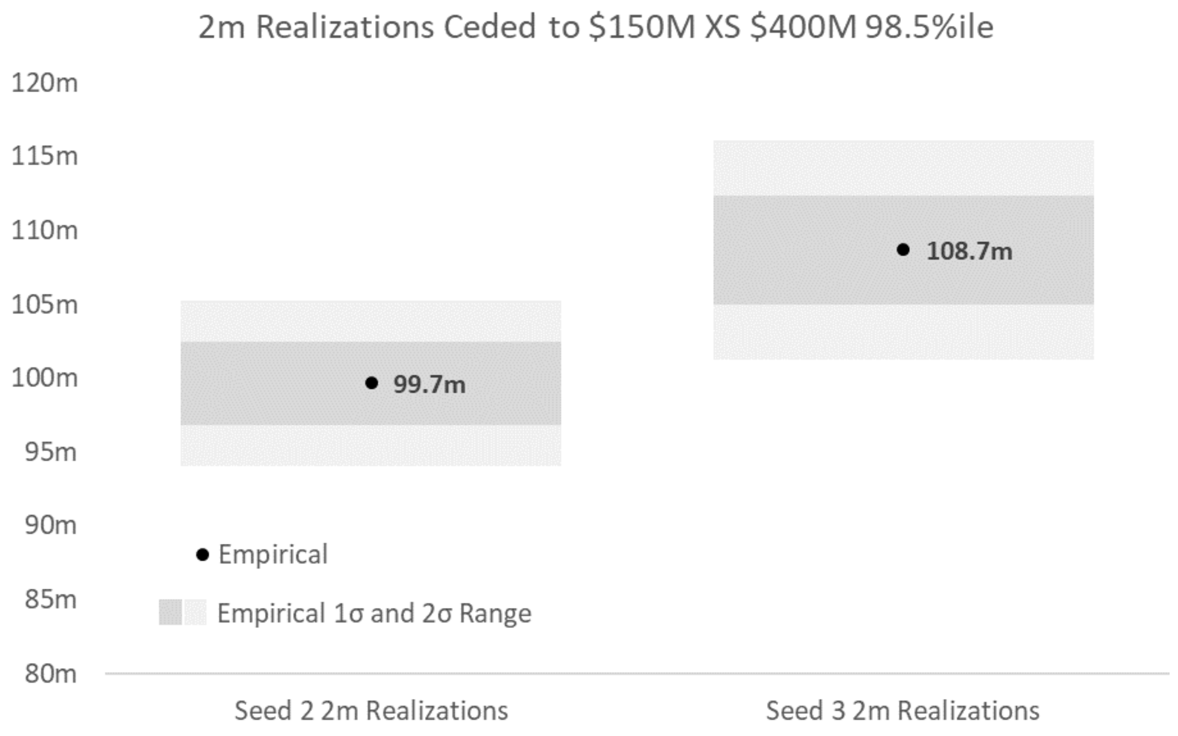


# How Reliable are Our Percentiles?

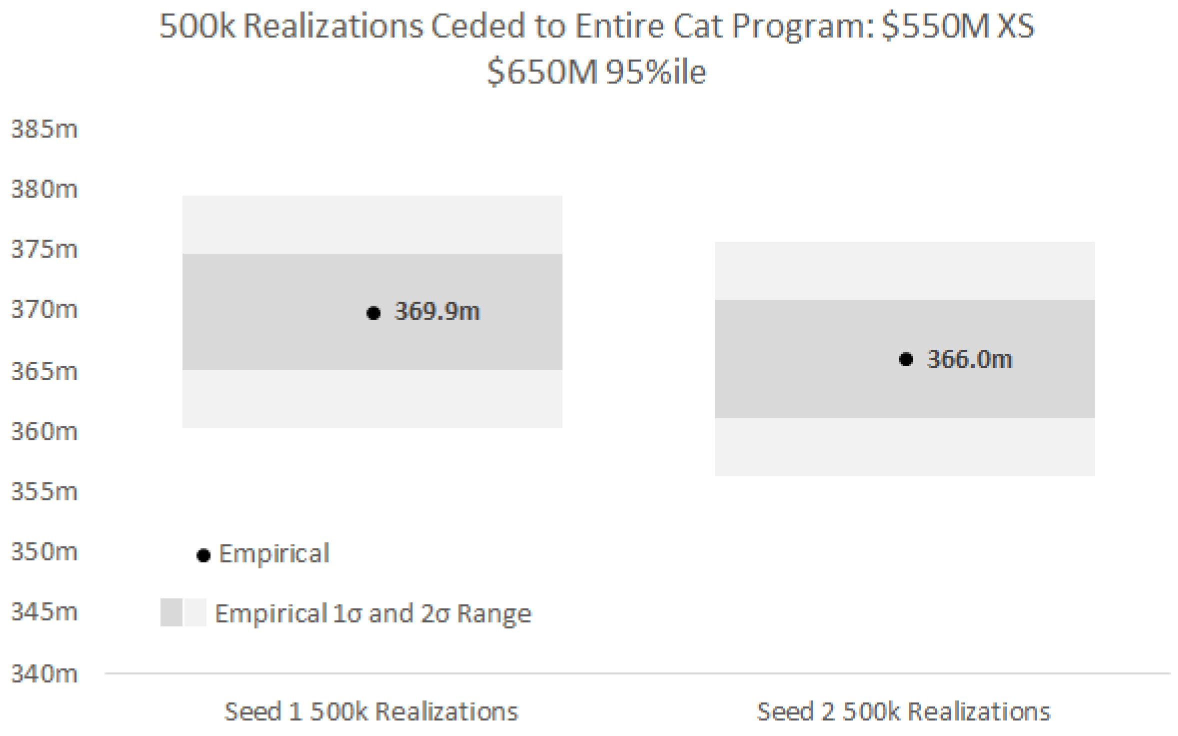
## High Excess Cat Using RMS is Problematic: Client 2

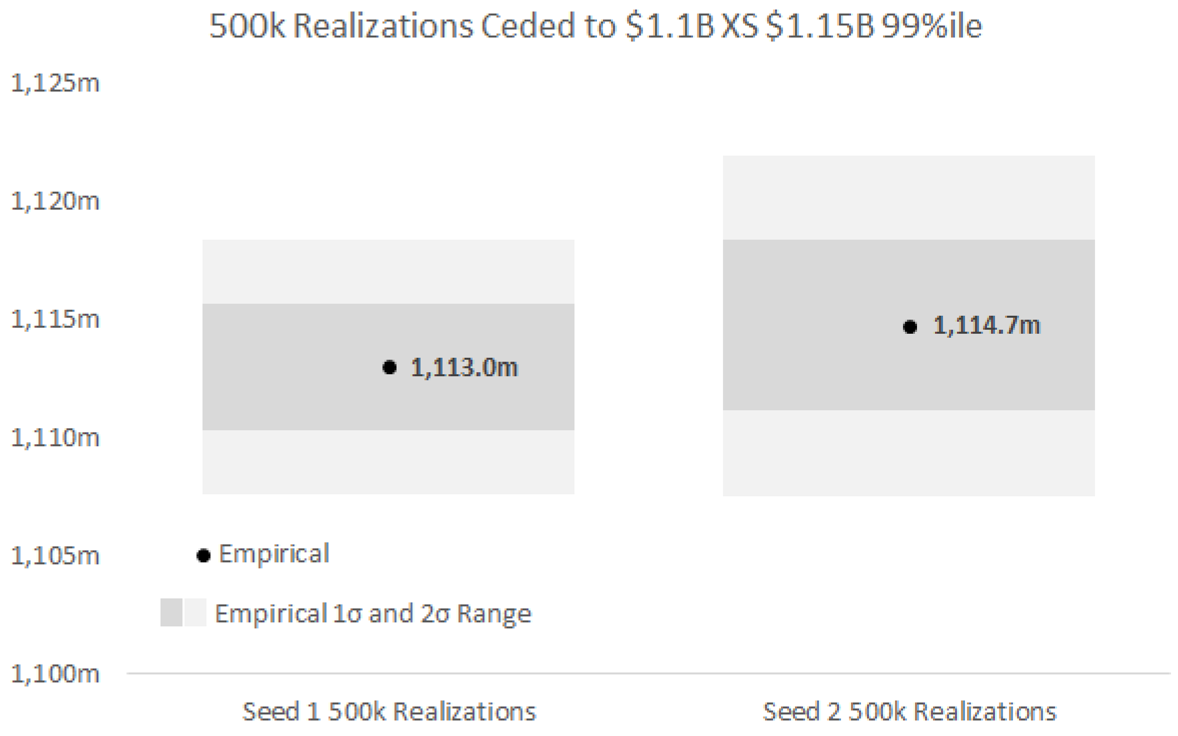


# How Reliable are Our Percentiles?

## High Excess Cat Using RMS is Problematic: Clients 3 and 4



# 6. Practical Implications and Recommendations

# Practical Implications and Recommendations

Ideally we would just simulate enough realizations so virtually all of our tail statistics more-or-less converge, but this seems impossible for our cat models in the sub-10m simulation regime

The estimated CV of any simulated percentile *itself* varies on re-simulation with CV $\frac{1}{\sqrt{Observations\ in\ Percentile\ Neighborhood}}$, so we should aim for $Obs > 100$

If your cat reinsurance decision is driven by RMS tail statistics and isn't diversified by region/peril then it is worthwhile simulating > 2m realizations

# Practical Implications and Recommendations

Ideally we would just simulate enough realizations so virtually all of our tail statistics more-or-less converge, but this seems impossible for our cat models in the sub-10m simulation regime

The estimated CV of any simulated percentile *itself* varies on re-simulation with CV $\frac{1}{\sqrt{Observations\ in\ Percentile\ Neighborhood}}$, so we should aim for $Obs\ >\ 100$

If your cat reinsurance decision is driven by RMS tail statistics and isn't diversified by region/peril then it is worthwhile simulating > 2m realizations

# Practical Implications and Recommendations

Ideally we would just simulate enough realizations so virtually all of our tail statistics more-or-less converge, but this seems impossible for our cat models in the sub-10m simulation regime

The estimated CV of any simulated percentile *itself* varies on re-simulation with CV $\frac{1}{\sqrt{Observations\ in\ Percentile\ Neighborhood}}$, so we should aim for $Obs\ >\ 100$

If your cat reinsurance decision is driven by RMS tail statistics and isn’t diversified by region/peril then it is worthwhile simulating > 2m realizations

# Open Problems

## Areas for Further Investigation and Improvement

- **Recommended Realizations:** More robust justification/s; including understanding the impact of book size, retention/limit, etc.

- **Irreducibly Empirical?** Is it even *possible* to make theory-grounded recommendations given the asymptotic results, our convoluted distributions, highly varied contracts & exposures, etc.? Maybe a deep theoretical understanding of RMS's end-to-end modeling could be used to infer *something*?

- **AIR Implications**: Better understanding of the implications for AIR-based models that include non-cat but for which we run 10k realizations to perfectly exhaust the 'catalog' – e.g., does the *a priori* 0 simulation error for the cat component produce low overall simulation error?

- **New Generation RMS:** Will upcoming RMS releases change our recommendations – e.g., if they shift to a kind of 'catalog' like AIR

- **Other Distributions and Contracts**: More work and experiments on other theoretical distributions, and on casualty cat / cyber

# 7. Appendices: Proof Outline and Further Details

# Appendices: Proof Outline and Further Details

## 1. Proof Outline and Thought Process

$\widehat{CV(\widehat{X_p})} \equiv \widehat{CV}$ using the neighborhood of O observations in the neighborhood around the percentile:

$$(1)\ \widehat{CV} \to \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}}-\widehat{X_{low}}}{\hat{X}}$$

$$(2)\ CV\left(\widehat{CV}\right) \to \frac{1}{\sqrt{O}}$$

Proof of (1) is simple using the asymptotic formula for the standard deviation, applying the inverse function derivative rule, and estimating the gradient as the gradient of the line segment connecting the left-most point of the neighborhood to the right-most point

Proof of (2) relies on a handful of theorems, properties and techniques, but to give some intuition:

- Simulations can be thought of as coming from U(0,1) and then just taking the corresponding x when mapped onto the CDF of the distribution of interest
- $CV\left(\widehat{CV}\right) \to CV\left(Range\colon \widehat{X_{high}} - \widehat{X_{low}}\right)$ because the Multiplier is a constant and $\frac{1}{\widehat{X_p}}$ is asymptotically independent of the gaps (per section 3.2 of 2017) and approaches a constant using the Delta Method or Slutsky's Theorem… So we can just focus on the CV of Range
- For the Uniform, it is easily shown that $CV\left(Range\colon \widehat{X_{high}} - \widehat{X_{low}}\right) = \sqrt{\frac{Samples+1-Obs\ in\ Hood}{Obs\ in\ Hood \times (Samples+2)}}$, which clearly approaches $\frac{1}{\sqrt{O}}$. The proof follows straightforwardly from the fact that the Range is Beta where X is Uniform
- For other distributions, here are two possible intuitions before giving the rigorous proof:
  1. The uncertainty (CV) in Range for any distribution is explained fully by the uncertainty (CV) in Range for the Uniform because of the first bullet in this section
  2. In the neighborhood of the pth percentile, the gap until the next observation can be thought of as n samples of your RV all constantly 'trying' at the same time to hit a value around the neighborhood, each having $f(x)$ probability, and us waiting until one hits -- this is essentially an Exponential distribution with parameter $\frac{1}{nf(x_p)}$ by the scaling property. E.g. if n is 4 and f(x) is a half then it's a wait time with expected value 2. Given that the gaps are asymptotically independent, Range is just the sum of these Exponentials, which has CV $\frac{1}{\sqrt{O}}$

The rigorous proof of (2) takes Theorem 1 from 2017 which runs roughly like my 2 above, and then the sum of these Exponentials is an Erlang that produces the CV of $\frac{1}{\sqrt{O}}$. After writing all this, I found out that 2020 S2.1 notes historical framings that are similar to mine but stop at the σ rather than the CV

# Appendices: Proof Outline and Further Details

## 2. Second-Order Uncertainty: Volatility of Percentile Estimates for Unknown Distributions

Instead of aiming for σ, we'll aim for the $CV = \frac{\sigma}{\mu}$ using the neighborhood of O observations around the percentile:

$$\widehat{CV} \to \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}} - \widehat{X_{low}}}{\hat{X}} = Multiplier \times Uncertainty\ \%\ in\ Obs\ Region$$

The Multiplier

- Increases as N increases, because zooming in shrinks the domain over which the gradient is calculated
- Decreases as the number of Observations in the neighborhood increases, reflecting more credible data/smoothing
- Is highest in the middle of the distribution because *#samples ≤ pth* percentile is *Bi(N,p)*

The Uncertainty %

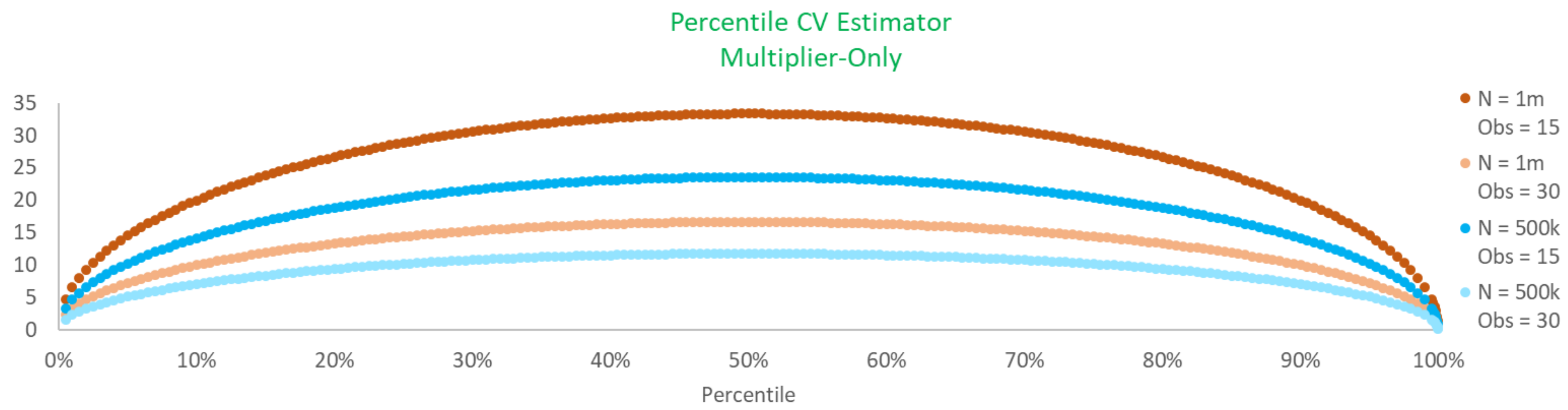
- Increases if the neighborhood is more volatile, which is partially caused by adding Observations to it (all else equal)
- Decreases if the neighborhood is less volatile, which is partially caused by increasing the number of realizations in the simulation so the neighborhood range clusters tighter with less uncertainty
- Is highest at sparse extreme regions and very low in the middle of the distribution (usually)

# Appendices: Proof Outline and Further Details

## 3i. Second-Order Uncertainty: Volatility of Percentile Estimates for some known Distributions

Instead of aiming for σ, we'll aim for the $CV = \frac{\sigma}{\mu}$ using the neighborhood of O observations around the percentile:

$$\widehat{CV} \rightarrow \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}} - \widehat{X_{low}}}{\hat{X}} = Multiplier \times Uncertainty\ \%\ in\ Obs\ Region$$

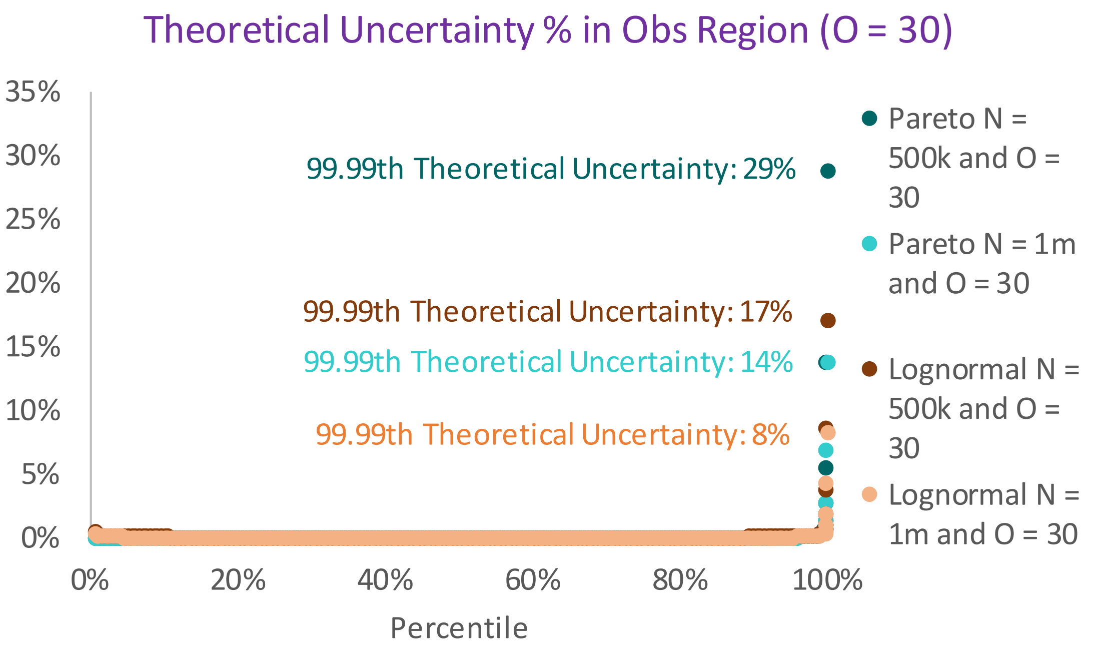


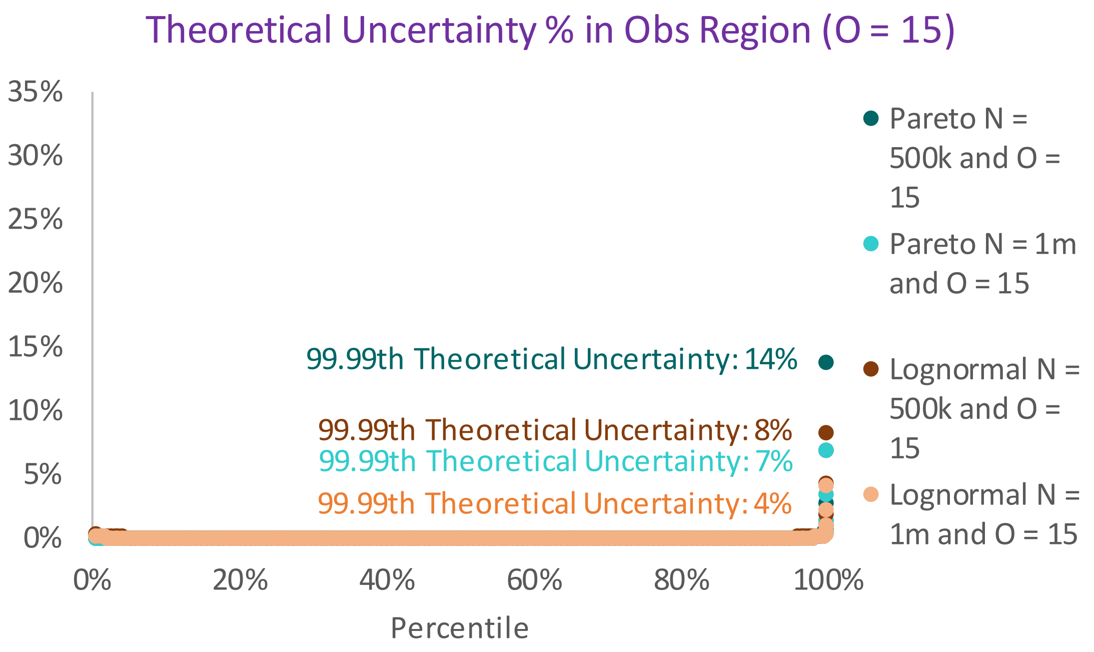
# Appendices: Proof Outline and Further Details

## 3ii. Second-Order Uncertainty: Volatility of Percentile Estimates for some known Distributions

Instead of aiming for σ, we'll aim for the $CV = \frac{\sigma}{\mu}$ using the neighborhood of O observations around the percentile:

$$\widehat{CV} \rightarrow \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}} - \widehat{X_{low}}}{\hat{X}} = Multiplier \times Uncertainty\ \%\ in\ Obs\ Region$$



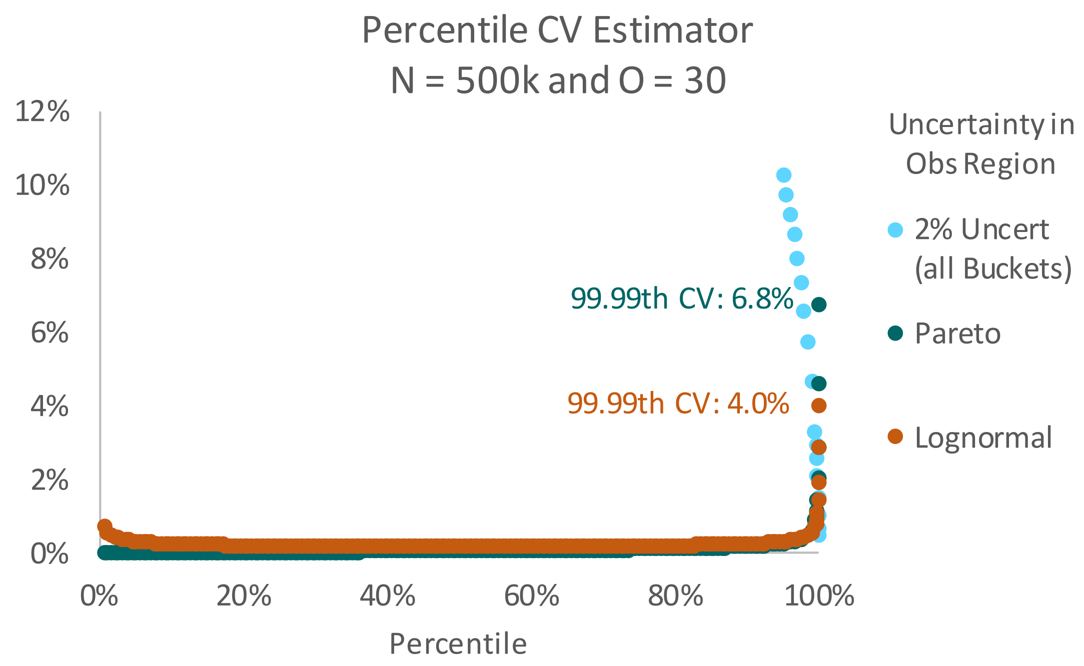



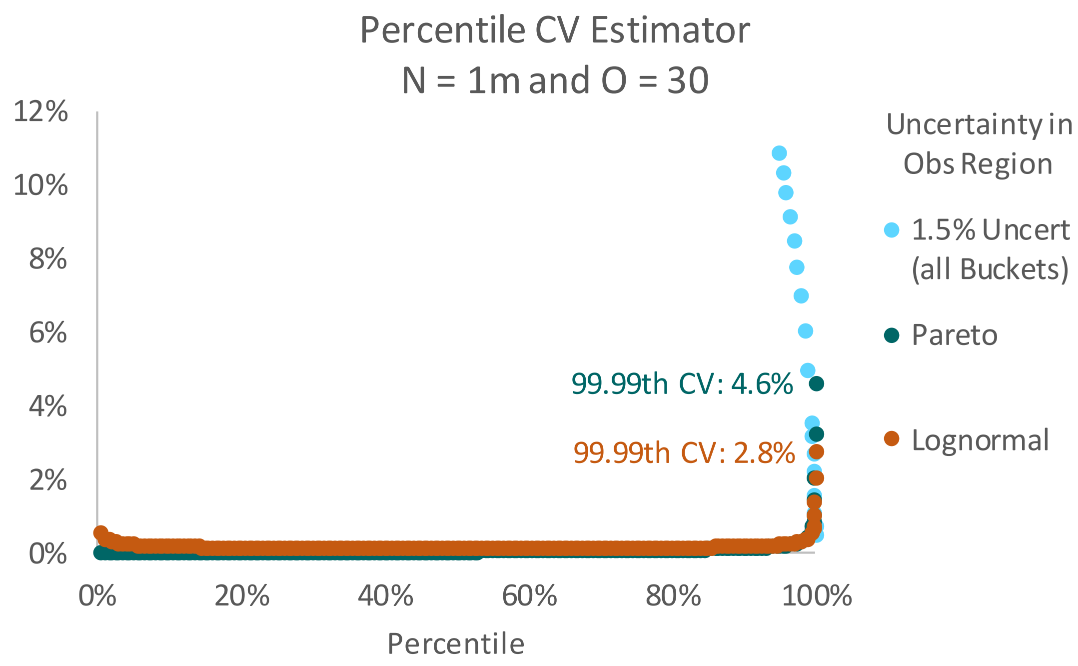
# Appendices: Proof Outline and Further Details

## 3iii. Second-Order Uncertainty: Volatility of Percentile Estimates for some known Distributions

Instead of aiming for σ, we'll aim for the $CV = \frac{\sigma}{\mu}$ using the neighborhood of O observations around the percentile:

$$\widehat{CV} \rightarrow \frac{\sqrt{Np(1-p)}}{O} \times \frac{\widehat{X_{high}} - \widehat{X_{low}}}{\hat{X}} = Multiplier \times Uncertainty\ \%\ in\ Obs\ Region$$



# Appendices: Proof Outline and Further Details

## 4i. Uncertainty Frameworks: Contextualizing Simulation Error

So we can think of simulation uncertainty as having two levels:

1. **First-Order Uncertainty**: Any simulated metric (random variable) like gross losses, net losses, ceded to a contract, etc. has percentiles – i.e., **a CDF**
2. **Second-Order Uncertainty**: The simulated percentiles *themselves* have inherent uncertainty and will change if we re-simulate – i.e., *our attempt at* **Error Bars** using the imperfect information from our simulation

And this can be contextualized within the following useful Error/Uncertainty decomposition:

- **Process/Stochastic Error**: The best *we* can do is model probabilities – we almost never predict metrics with certainty, even if we have the best data, model, and parameters; and we use simulations to represent the randomness
- **Parameter Error**: Even if our model structure and logic are perfect, our parameter selections are limited by the quality, quantity and projectibility of historical data & our parameter selection procedure/s
- **Model Error**: Even if we've perfectly selected parameters, the model inputs, structure and logic might not perfectly represent the dynamics of the thing being modeled
- **Data and Assumed Constants Error**: There may be errors in the data provided, or in the assumed constants (like inflation) that we use

# Appendices: Proof Outline and Further Details

## 4ii. Uncertainty Frameworks: Broader Considerations and Value-Add

The present focus on Simulation Uncertainty ignores whether our data and model are any good to begin with!

Minimizing Simulation Uncertainty is like carefully checking the spelling and grammar on an email

- It can't fix the underlying reasoning, content, structure, messaging, style, etc.
- It *occasionally* reveals substantive issues, like when a typo or grammatical error changes the meaning of a sentence
- It shouldn't be the case that one of the drivers of results/YOYs is the random seed that we happened to use

Think carefully about the other sources of uncertainty

- Data Error: Extensive data validation tests, YOY comparisons, outlier detections, system upgrades/reviews, etc.
- Parameter Error: Sensitivity tests on key parameters, bootstrapping, parameter uncertainty in vendor models, etc.
- Model Error: Alternative selection methodologies, model structures & types, levels of granularity, etc.

GC's value prop is essentially to help clients navigate uncertainty

- Provide sensitivity tests, uncertainty bands, appropriate caveats/limitation, and rounded results
- Consider Cat Model adjustments – speak to the client and GC Model Evaluation
- Use brokers and GC's market knowledge to inform uncertainty on execution risk, market capacity, pricing, etc.

GuyCarpenter

A business of Marsh McLennan