MPRA
Munich Personal RePEc Archive

Uniform Confidence/Certainty Estimation

Arnold, Rob

Acorn Pass

7 August 2023

Online at https://mpra.ub.uni-muenchen.de/118234/
MPRA Paper No. 118234, posted 09 Aug 2023 13:30 UTC

Uniform Confidence/Certainty Estimation

Aug 27, 2023

V1.2

Rob Arnold - Acorn Pass, LLC
AcornPass.com

CC- BY-SA

Abstract

Uniform Confidence/Certainty Estimation (UC2) is an approach and set of tools that address several issues that are common in risk estimation techniques. Deployed between analysis and modeling, UC2 brings uniformity and interoperability that improve risk model results and improve stakeholder engagement. Its unique features correctly capture confidence and certainty and improve interoperability between data-driven and expert-derived risk estimates and the models that consume them. In turn, UC2 increases uniformity, transparency, and stakeholder engagement, without ripping and replacing existing risk models or analytical workflows.

Keywords: Risk modeling, Risk analysis, Risk estimation, Scales, Confidence, Certainty, Accuracy and Precision, Quantitative, Qualitative, Objective, Subjective, Binomial probability, Probability distribution.

Table of Contents

Abstract **0**
Table of Contents **1**
Introduction **2**
About the Author 4
Risk Estimation Issues **4**
Data Types 5
Data Origin 5
Resource/Data Tradeoffs 6
Estimate Aggregation 7
Existing Techniques **8**
ICD 203 Scales 8
PERT 10
UC2 Scales **11**
Assured Range and Segments 12
Unknown 12
Confidence Bands 12
UC2 Field Usage **12**
Target Risk Resolution 12
Unknown 13
Confidence Bands 14
Gather Estimations 15
UC2 Analysis **16**
Interpreting UC2 Scales 16
UC2 Distributions 17
Risk Model Integration 18
Binomial Inputs 18
PERT Inputs 18
Distribution Inputs 18
UC2 Risk Resolution 19
Conclusion **19**
Appendix A - ICD 203 Python Code **21**
Appendix B - UC2 Qualitative Use Case **23**
Appendix C - UC2 Color Pallets **26**

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

About the Author

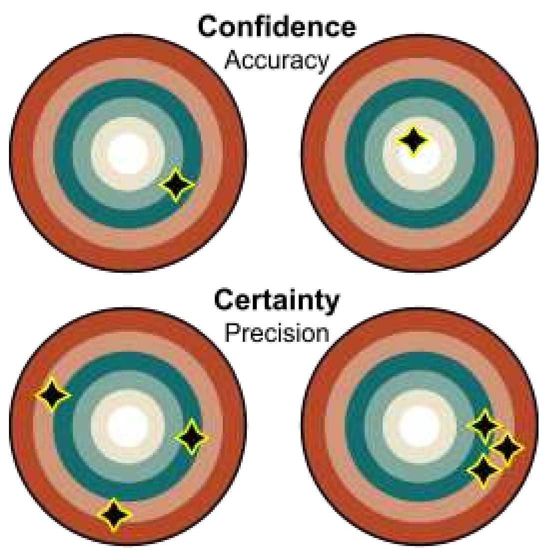
I had the honor of practicing risk management at levels ranging from the smallest, and most resource constrained organization to national and global risk. It started with cyber risk management focused squarely on underserved small businesses, nonprofits and local government organizations. From there it evolved into include the COVID-19 pandemic, national security, and global supply chain risks. My expertise has been lent to the National Institute for Standards and Technology (NIST), the Department of Homeland Security (DHS), the Cybersecurity and Infrastructure Security Agency (CISA), multiple Whitehouse cabinet members, and members of Congress. UC2 is the culmination of a decade's worth of academic and practical risk management that is being given back to the risk management communities which afforded me these diverse opportunities.

Introduction

At the center of risk analysis is the issue of confidence. This problem is widely acknowledged, but often misunderstood and not well managed.

Confidence refers to the ability of an expert to make subjective, accurate predictions that align with the objective truth. If truth is the bullseye of a target, confidence represents how close risk estimations are to the center. Confidence is the *expected* proximity to truth.

Certainty is the agreement between multiple estimates from many sources. Tightly grouped estimates indicate more certainty; more dispersion means less. Certainty is a measure of precision and consistency, not closeness to truth.



This paper introduces Uniform Confidence/Certainty Estimation (UC2 / `you-see-two`) as a solution to this and other problems in estimating values for use in risk equations and risk models. It is applicable to quantitative, data-driven estimation techniques as well as qualitative techniques. It also acts as a bridge between quantitative, qualitative, objective, and subjective estimations freeing the risk analyst to use most combinations of data sources together in a uniform and transparent manner.

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

UC2 Scale - Quantitative Example

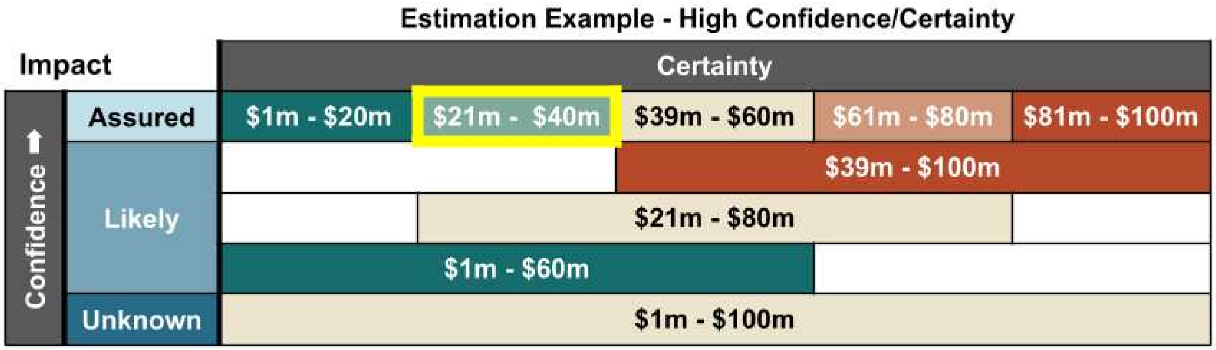
Usage by risk analysts and subject matter experts (experts) is straightforward and intuitive. Segments across the top row of the UC2 Scale express the *desired* level of granularity for an estimate of an arbitrary range. This is the bullseye of objective truth with just enough granularity to assist a decision making stakeholder — or risk model — arrive at an *actionable* outcome.

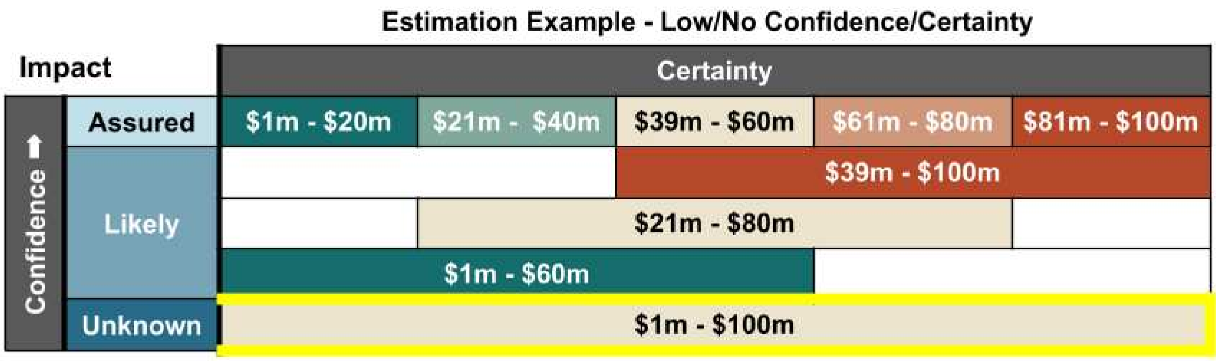
When data-driven or expert-derived estimates cannot be neatly mapped to the bullseye, the target confidence/certainty drops to lower levels. This example also illustrates overlapping certainties, which are common in both data-driven and expert-derived estimation.

The bottom row, labeled “Unknown” is the very lowest level of confidence. In data-driven estimations it denotes that an acceptable estimate was not found within the data. Experts will use this row to indicate “I don’t know” or “I lack context” In both cases it can also indicate that no data or expert source is available.

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

Uniform Confidence/Certainty Estimation (UC2)

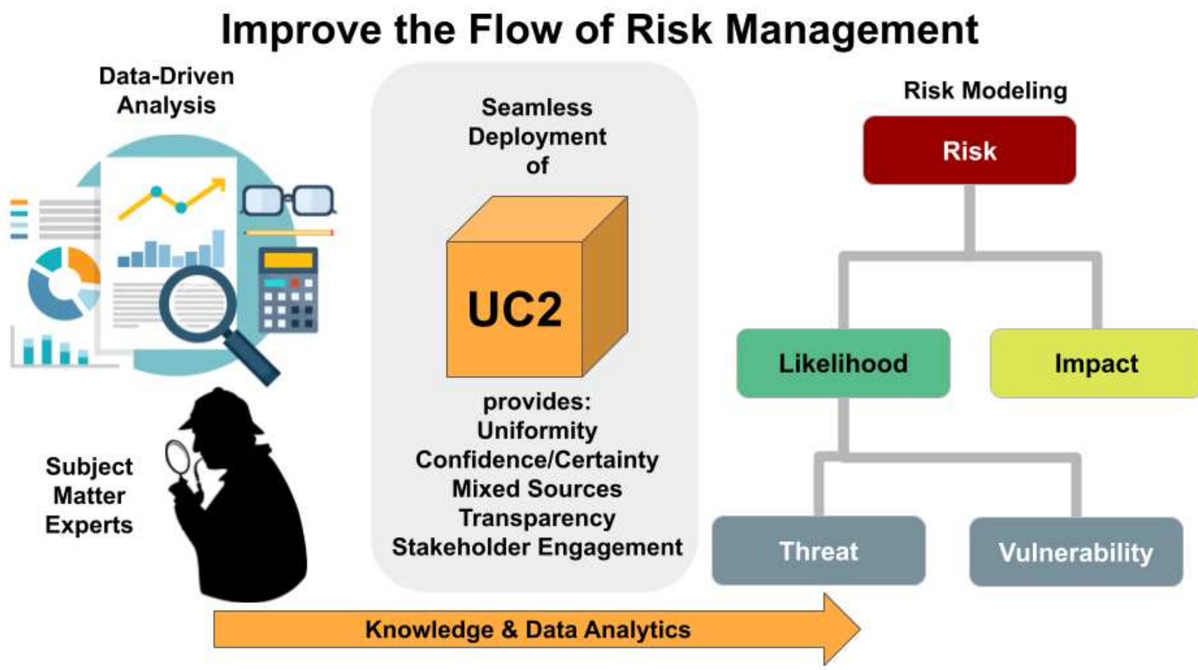




Mapping data-driven and expert estimations to UC2 explicitly captures both confidence and certainty in a uniform manner that allows for aggregating estimates from both data-driven and expert sources. As further discussed in “UC2 Analysis”, multiple estimates combine in a way that transparently honors confidence and certainty. UC2 Distributions are more nuanced and accurate outputs that are compatible with nearly any risk model or risk assessment.

With respect to integration with existing models and workflows, UC2 offers incremental improvement *without* having to rip and replace existing models and estimation techniques. It fits seamlessly into existing data- and expert-driven workflows making them more uniformly compatible through UC2 Scales and UC2 Analysis. UC2 Distributions integrate just as seamlessly with nearly any model of risk. The outputs are compatible with *existing* model inputs like binomials, traditional PERT estimates, PERT distributions, and other free-form distributions.

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com



Risk Estimation Issues

Confidence and certainty are not the only issues plaguing risk estimation. This section describes additional problems starting with issues that arise from conflating data types and data origins. The tradeoff between precision and data management is discussed through the lens of Risk Resolution. Issues surrounding default values and expert biases round out the topics in this section.

Data Types

Quantitative: An expression of quantity using numbers where the meanings and proportionality of values are maintained inside and outside the context of the assessment[1].

Qualitative: An expression of the nature of something based on non-numerical categories, levels[2] or narrative description.

Ideally, all elements of risk would be reducible to a quantitative measure that was independent in any context. The reality is that nearly every risk analysis relies upon a mixture of the two.

[1] DHS Risk Lexicon 2010 Edition (Quantitative Risk Assessment Methodology)
[2] DHS Risk Lexicon 2010 Edition (Qualitative Risk Assessment Methodology)

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

Data Origin

Objective: An unbiased, impersonal observation of something.

Subjective: An interpretation or estimate of probability as a personal judgment or degree of belief about how likely a particular event is to occur, based on the state of knowledge and available evidence[3]. Subjectivity extends beyond likelihood to other model inputs such as impact.

Objectivity and subjectivity are *independent* concepts from quantitative or qualitative data types. In particular, a harmful idea lies in thinking that quantitative data is always objective. In practice, quantitative risk management data is frequently derived from very subjective sources.

Objective data from actual observations has a very high degree of confidence. The certainty may be all over the board, but the observations are about as faithful to truth as one could hope. Data derived from experts that spent a lifetime observing relevant scenarios, can be just as confident and highly certain in some cases.

Quantitative data is often favored because it has a very measurable certainty which makes it appealing to under-trained risk analysis. The result is that certainty masquerading as confidence has dire consequences for unsuspecting stakeholders because closeness to truth far outweighs data consistency when making critical decisions. Presenting certainty as confidence exacerbates this issue.

Problem One: *Capturing confidence at the origin, regardless of the type of data.*

Another issue with objectively sourced, quantitative data is that objectivity is extremely contextual. Taken out of the context from which it originated, it becomes subjective and takes on a different level of confidence as a result. Using objective lightning strike data collected in Colorado to predict lightning strikes in North Carolina is an example of taking high confidence, Objective data from one context and applying it subjectively in another context. The issue here is that the researcher's own confidence that Colorado data is close to North Carolina's truth changes the overall confidence.

Problem Two: *Confidence in objective data can change at the point of application due to a Subjective change in context.*

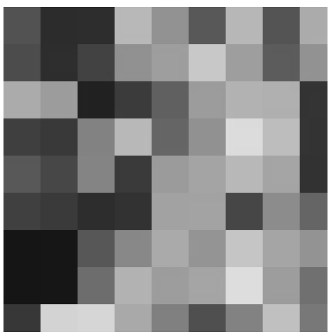
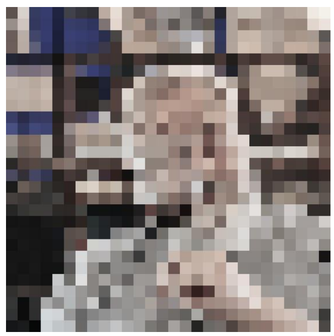
[3] DHS Risk Lexicon 2010 Edition (Subjective Probability)

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

These problems, grounded in misunderstandings of key concepts, play out in very subtle ways in the real world. As described in the "Existing Techniques" section, these problems can survive for decades even in well-run risk management programs.

Resource/Data Tradeoffs

Risk Resolution — an image-resolution metaphor — is an excellent way to frame the demand for a higher resolution view of the risk landscape juxtaposed with the problem of sourcing and storing an infinitely large mountain of risk data. With only a handful of data "pixels" defining a given risk, the clarity will be relatively low. Increasing the pixel count and color-complexity drives up data-related costs.

Risk Resolution		
		
Small Data	**Big Data**	**Extreme Data**

Risk Resolution is also a way to explain this problem to stakeholders. It should be noted that the quest to fill in missing data leads risk analysts to apply objective data in subjective contexts, without capturing the change in confidence. As a result, stakeholders are given a vivid picture of the risk landscape composed of tainted data.

Problem Three: Increasing Risk Resolution requires enormous amounts of data.

Estimate Aggregation

In situations such as the rapidly unfolding COVID-19 pandemic for which I was hired to help advise CISA's National Risk Management Center[4], estimates of risk were arriving in every form one could imagine. Sometimes estimates were quantitative, sometimes they were qualitative and never with a uniform measure of confidence. It was a nightmare to match estimates given in one scale with those given in another.

[4] https://www.cisa.gov/about/divisions-offices/national-risk-management-center

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

The most common approach is to use only qualitative or only quantitative data to estimate a single value. While that neatly avoids the problems associated with combining them, what happens when new data sources emerge that use the other? Transitioning a risk model from qualitative data to quantitative data represents a conundrum. Typically, either the old data is tossed out or it is somehow translated into quantitative terms at significant cost in time and effort.

Matching estimates across scales is hard when each source defines their own scale. Even when two scales have the same range, the segments might differ. There is also the conundrum of expanding scales. When COVID-19 became much worse than initially expected, this particular issue reared its head.

Problem Four: Incompatibility between data types and non-uniform scales.

These problems are by no means an exhaustive list of risk analysis issues, but they represent a pretty wide swath of the fundamental problems many risk management programs face. They plague new risk management programs run by newcomers to the field as well as established programs run by seasoned risk managers.

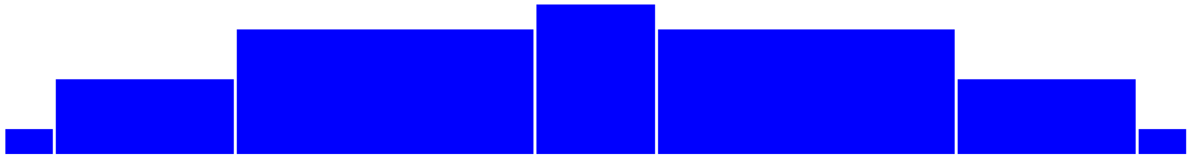
Existing Techniques

This section examines two real-world techniques for capturing estimates of risk. These techniques represent the issues faced by data-driven and expert-derived estimation. For example, ICD 203 Scales have numeric definitions that would be targets for a Bayesian natural language classifier. Meanwhile, PERT is representative of data-driven techniques based on descriptive statistics such as range, median values, and averages.

ICD 203 Scales

Below a visual is used to examine the structure and *intent* of a scale. As it is defined by the Intelligence Community, a predetermined “shape” can clearly be seen in the scale’s quantitative definition. The columns shown above the ICD 203 likelihood scale have their widths set in proportion to the ICD 203 definition; this helps visualize the scales’s structure. The pattern of widths corresponds to the typical segmentation of a normal distribution — the bell shaped curve. Column height is used to illustrate the *intended* (but not actual) effect of building a normally distributed scale.

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com



almost no chance	very unlikely	unlikely	roughly even chance	likely	very likely	almost certain(ly)
remote	highly improbable	improbable (improbably)	roughly even odds	probable (probably)	highly probable	nearly certain
01-05%	05-20%	20-45%	45-55%	55-80%	80-95%	95-99%

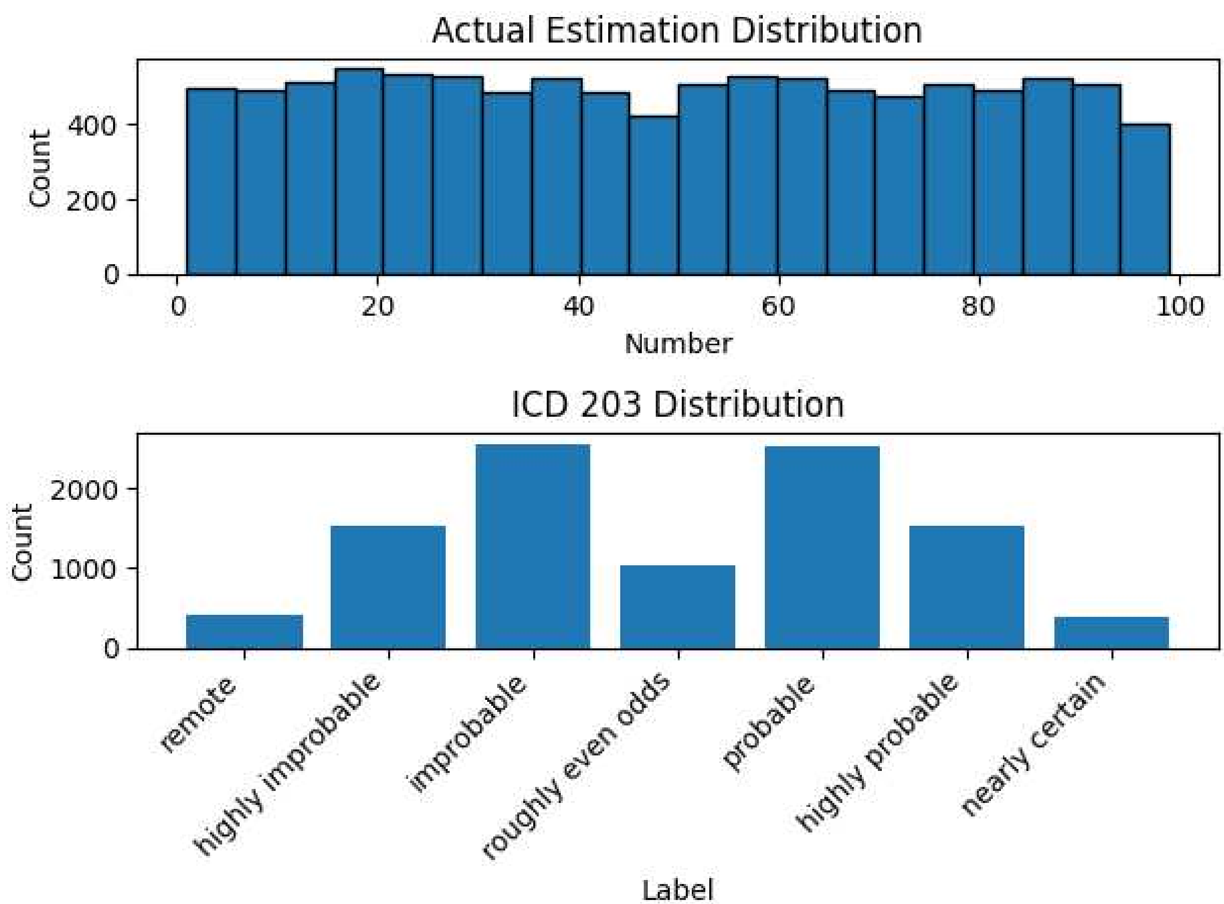
Table Source: 2015 Intelligence Community Directive (ICD) 203 - https://www.dni.gov/files/documents/ICD/ICD%20203%20Analytic%20Standards.pdf

The segments of the scale are labeled with a mixture of terms relating to both certainty and confidence. With no further explanation in the directive, the analyst is left to decide on a case-by-case basis if they are going to report confidence or certainty. These are two different concepts that are not interchangeable. Upon aggregation of many different estimates, obvious problems will result from the conflation of the two concepts. This is symptomatic of problem one.

Well meaning intelligence experts appear to have designed this scale with the notion that estimates should fall into a normal distribution. This is clearly an attempt to coerce an assumed-by-the-scale-author level of confidence or certainty into the independent estimates of intelligence analysts. The goal was to add rigor, but the results illustrate problem two — additional subjective confidence added at application.

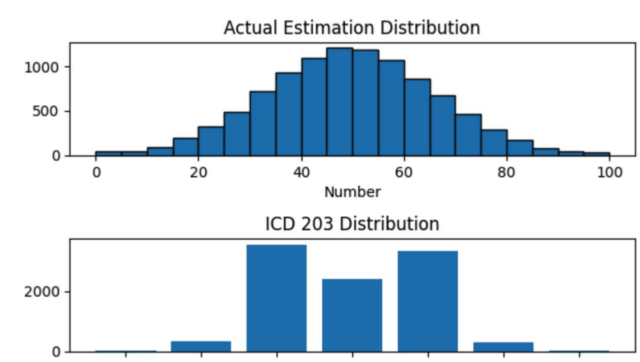
Let's look at an example. Take the case where no one really knows the answer, thus all likelihoods are equal. In this case, estimations would fall into a uniform distribution, which is flat and extends from the minimum to the maximum. Below is an example built on 10,000 estimates between 1 and 99% representing an unknown likelihood The first graph shows the distribution of estimates. The second graph maps the same estimates to their ICD 203 categories. The shape of the data going in gets skewed and perhaps not as one might expect!

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com



Despite being built from the exact same set of data, the two graphs look nothing alike and interpreting the data through the ICD 203 lens is more complicated. With the raw data it is clear that all likelihoods are equally possible. ICD 203 erroneously suggests otherwise.

To the right the process is repeated with a normal distribution of estimates, the ICD 203 results retain the same odd shape as before. The scale's segmentation has a big impact on the results of the data. It is almost certain that the ICD 203 authors did not intend this shape.



See Appendix A and try it yourself in a Python notebook.

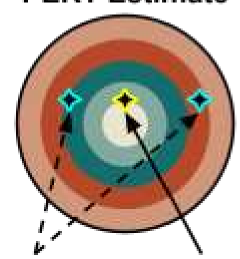
CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

These graphs clearly illustrate how improper scales and unclear measures of certainty and confidence lead to major issues in aggregated analysis and communication to stakeholders.

There are use-cases for variable-length scales that can help researchers correctly capture and analyze accuracy and confidence alongside value estimates. The Intelligence Community had the right idea; the implementation was simply imperfect. A UC2 Scale could achieve what the authors of ICD 203 were hoping to accomplish. It uses variable width scales correctly to allow the analyst — not the researcher — to express a level of confidence in their estimate that aids rather than complicates analysis.

PERT

Program Evaluation Review Technique[5] (PERT) is a widely used technique in risk management to derive values from expert estimates for quantitative model inputs when objective data is unavailable. PERT and its derivatives surely have their place in the risk analyst's toolkit, but the caveats should be well understood.



When giving an estimate, experts often choose to give ranges rather than a single value. PERT and its derivatives explicitly ask for this as well as a most likely value. The width of the range can be seen as an estimate of certainty and the most likely value is an attempt to confidently predict the correct value. PERT distills these three estimates to a single value through a weighted average. The argument for PERT-based methods is that the resulting value captures the expert's confidence as the most likely value that is then adjusted by the range for certainty.

The good thing is that PERT independently collects a measure of confidence and certainty — that's great! But then PERT makes a weighted average from the estimates per its definition which gives the most likely value four times the weight of the other values.

$$PERT\ Estimate\ =\ \frac{(4 \bullet Most\ Likely) + Range^{Min} + Range^{Max}}{6}$$

[5] Engwall, Mats. (2012). PERT, Polaris, and the realities of project execution. International Journal of Managing Projects in Business. 5. 10.1108/17538371211268898.
See also: https://www.historicprojects.com/PERT.html

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

Once solved for the estimate, the information about confidence and certainty is discarded. This circles back to problem one — confidence is not collected or kept in this case. Beta-PERT estimates, which allow for weights other than 4, suffer the same problem of losing confidence and certainty.

The PERT distribution is another widely used variation. It is an excellent technique in most respects, especially in the domain of project management where it was born. This technique attempts to retain the original confidence information by interpreting a curve from the original PERT-estimates, then deploying the curve to the model as a distribution. The problem with PERT distributions is that a risk analyst — *who is not the original knowledge source* — is deciding which distribution pattern to apply uniform (bell shaped), normal (flat), etc. Retroactively adding information that was not directly expressed by the source leaves PERT distributions exposed to Problem Two — subjective confidence is introduced by risk analysts after data gathering. The UC2 Analysis section will add more clarity on this topic.

UC2 Scales

The name **UC2 Scale** is used to differentiate the visual UC2 estimation diagrams from the larger UC2 methodology. This section examines the main components of a UC2 Scale, which are tagged in the image below and described thereafter.

Assured Range and Segments 1

Together, the range and segments represent the bullseye for estimation with fully **Assured** confidence and certainty.

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

Unknown 2

The opposite of the Assured bullseye, is the **Unknown** scale. This represents missing data or inadequate information/knowledge to make an estimate at any level.

Confidence Bands 3

One or more optional UC2 **Confidence Bands** reside between the extremes of Unknown and Assured. Their labels connotate a position between Unknown to Assure. Segments in a Confidence Band represent the degree to which an estimate can be made with less confidence and certainty than Assured, but more than just Unknown.

UC2 Field Usage

The process of setting up UC2 Scales and using them to capture data-driven and expert-derived estimations of risk elements is **UC2 Field Usage**. With respect to problem four — interoperability — UC2 is designed to work equally well with data-driven or expert-driven estimates. So, the risk analyst is free to use either or both. This is explained in more detail during the discussion of UC2-based analysis, but the basic setup and mandatory elements are laid forth in this section.

Target Risk Resolution

A uniform, fixed-width Assured scale is a *mandatory* element of a UC2 Scale. The range might be determined based on the requirement of an existing risk model's input. In which case, segments are evenly distributed across that range with the segment count being one that makes sense in the context of the risk model.

When model guidance does not exist, the risk analyst works with decision making stakeholders to establish the bullseye for estimation in terms of range and segmentation at the Assured level. Establishing the range is pretty simple. Just set the highest and lowest bounds that matter to the decision at hand. The segment count is likewise established with respect to how much resolution is necessary within the range to make *actionable* decisions. Together the range and segments define the bullseye that data-driven or expert-derived estimates will attempt to hit.

As mentioned in the Introduction, UC2 Scales can be expanded later, so getting the range and segment count right for all-eternity is not an issue. The practical limit is ten segments, but the goal is to pick the least number of segments necessary to assist the stakeholder with the decision at hand. In theory, one could make a scale long enough to

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

capture any level of precision. In practice, this is impractical due to problem three — the tradeoff of Risk Resolution vs. data acquisition and analysis.

A lower segment count also acts as a guard against over-confidence with estimates. It also helps counter the Dunning-Kreuger effect with expert estimations. In data-driven usage, it protects against data scientists taking their analysis to the n^{th} degree; analysis-paralysis that wastes resources chasing unnecessary levels of granularity.

Unknown

Another *mandatory* element of a UC2 implementation is the Unknown scale. It is an incredibly simple scale that has *only* one segment that spans the entire range defined above. The Unknown scale is necessary for both quantitative and qualitative data types.

Unknown Scale

UC2 Scale		Certainty
Conf	Unknown	1% - 100%

Here are two minimalist examples of a UC2 Scale. In the first example risk estimates are captured at one of two levels of confidence: Assured or Unknown.

Minimal, Qualitative, Three-Point UC2 Scale (Likelihood or Impact)

UC2 Scale		Certainty		
Conf	Assured	High	Medium	Low
	Unknown	Medium		

Using the color and label from the middle segment as the color and label for the Unknown is deliberate. In this particular context “Medium” is meant to be a midpoint. Had the context called for a scale anchored at “Low” then low supplies the color and label, and so forth.

Minimal, Quantitative, Three-Point UC2 Scale (Likelihood)

UC2 Scale		Certainty		
Conf	Assured	1%- 33%	34% - 65%	67% - 100%
	Unknown	1% - 100%		

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

In quantitative contexts, the best choice for a segment label is the entire range. In certain situations, Unknown could be anchored at top or bottom of the range, the anchor lending its color and label to Unknown.

A data-driven estimation can map to either of these UC2 Scales using standard data analysis techniques to a particular segment on the Assured level. If that is not achievable due to poor or missing data, Unknown is chosen. expert estimates use it the same way: either they have enough knowledge and confidence to pick a segment at the Assured level, or they choose the sole segment in Unknown.

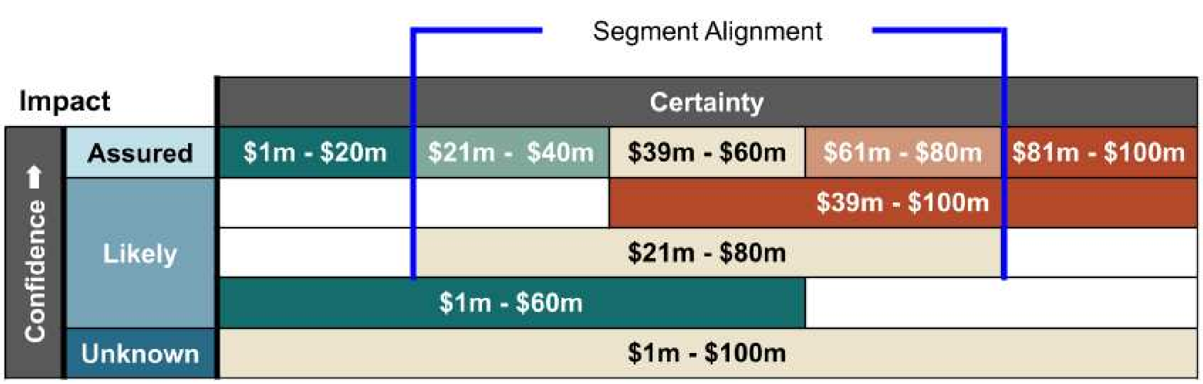
This *begins* to illustrate how UC2 addresses problem four — interoperability between data-driven or expert-driven data sources. There is more to that element of the solution in a later section.

Confidence Bands

Confidence Bands between the top row and Unknown is *optional*, but highly encouraged. When necessary, these intermediate bands can also have swim lanes within them to handle overlapping segments.

The number of Confidence Bands added to a UC2 Scale is left to the risk analyst and they can be added or removed at any point without running the risk of misaligning with previous estimation data. This feature is very handy when, in the middle of data-analysis or expert elicitation session, someone realizes that another level of confidence needs to be added on the fly.

When Confidence Bands are included, it is *mandatory* that the segments within be aligned to the boundaries of the Assured segments as illustrated below. This rule may not be intuitive for quantitative scenarios, but it exists to facilitate interoperability between quantitative and qualitative data sources.



CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

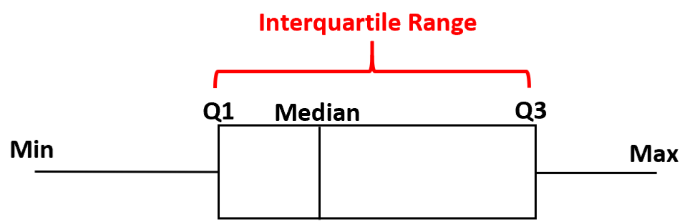
Confidence Bands are given labels that connote a closeness to the bullseye going from Unknown to Assured. For now, UC2 does not assign numerical values to Confidence Bands, which is an area for future research and development.

Gather Estimations

The risk analyst then uses one or more estimation techniques to select a *single* segment on the UC2 Scale. Examples techniques include:

- Scour reports, research and news for data that informs an estimate.
- Use data science to analyze machine readable datasets.
- Present representative risk scenarios to a subject matter expert along with the UC2 Scale and brief instruction on its use.

Regardless of the technique it is *mandatory* that an estimate must fit completely within a single segment defined by Assured, Unknown or any Confidence Band. This is fairly straightforward in expert-driven and qualitative scenarios.



In a data-driven scenario, descriptive statistics might summarize the contents of a data-stream as a box-plot where the interquartile range is taken to be the estimate. In this case the *entire* interquartile range must fit within a single segment. If adding a Confidence Band on the fly cannot resolve this, then Unknown should be selected on the UC2 Scale to represent that the data cannot represent an estimate within the confidence and certainty defined by the scale.

UC2 Analysis

After collecting estimates, **UC2 Analysis** is used to interpret the results and transform them into something that existing risk models and risk equations can easily ingest. The example given here is based on quantitative estimates. See Appendix B for an example using qualitative estimates.

Interpreting UC2 Scales

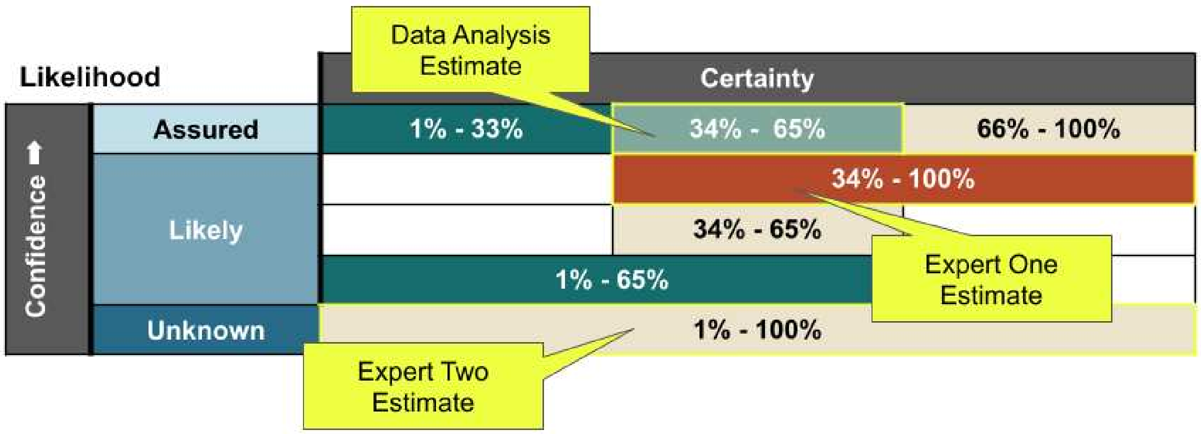
Each UC2 Scale-based estimate can be expressed as a small, atomic, uniform distribution. These become building blocks for an aggregate distribution that retains the

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

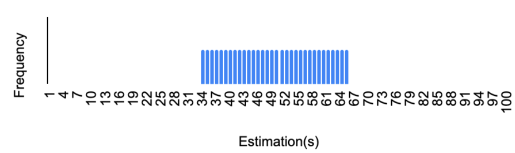
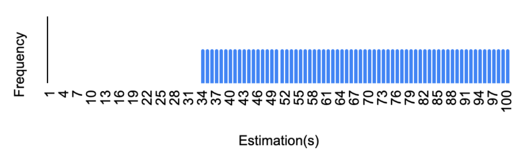
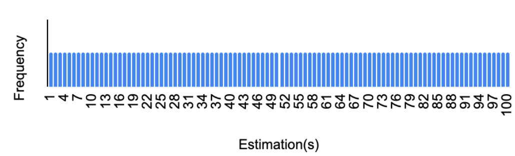
confidence and certainty characteristics imparted when individual estimates were forged at the source or subjectively applied as a proxy.

Metadata about which UC2 Scale was used for elicitation and the explicit level of confidence chosen should also be retained as they will be helpful later. Below is an example of three different Likelihood estimates on the same UC2 Scale. The table that follows illustrates the metadata and the atomic, uniform distributions extracted for each estimate.

UC2 Scale - Acme, Inc Likelihood v1.1



UC2 Estimation Examples

Meta Data			Interpretation	
UC2 Scale	**Source**	**Confidence**	**Estimate**	**Atomic Uniform Distribution**
Acme, Inc Likelihood v1.1	Data Analysis	Assured	34% - 65%	
Acme, Inc Likelihood v1.1	Expert One	Likely	66% - 100%	
Acme, Inc Likelihood v1.1	Expert Two	Unknown	0% - 100%	

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

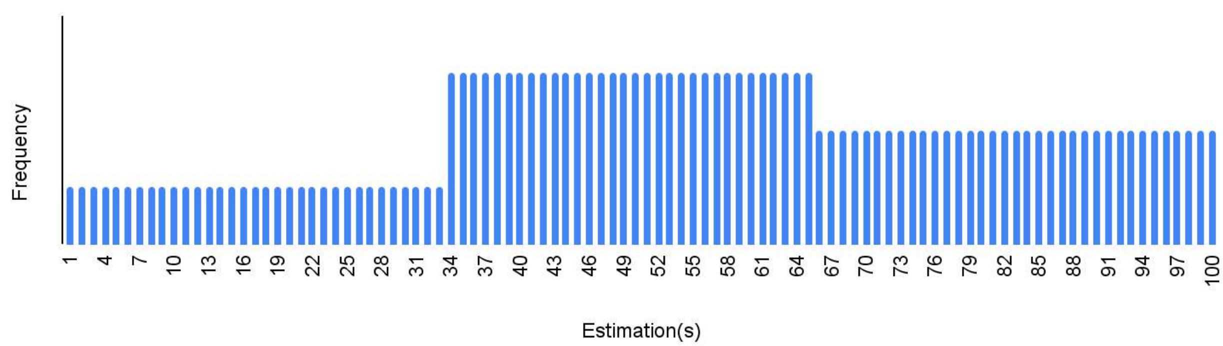
For brevity, the data label “Acme, Inc Likelihood v1.1” is a pointer to UC2 Scale parameters like the range, then number of segments defined for Assured, the anchor point for Unknown, the makeup of Confidence Bands and their relative position between Assured and Unknown. Again, this metadata will be useful later on, but the next step is carried out without this additional information.

UC2 Distributions

Building distributions is far better than averaging. A full discussion is outside the scope of this paper, but Sam Savage’s book “The Flaw of Averages” makes the case clear[6]. So, the next step is to build composite distributions from the interpretation step.

UC2 Distributions are built by stacking the atomic, uniform distributions derived in step one. The graph below illustrates the UC2 Distribution based on the likelihood estimates provided above.

UC2 Distribution



The resulting UC2 Distribution conveys confidence and certainty as intrinsic elements of its structure. Unlike PERT distributions which take on a shape defined by someone *after* the elicitation, UC2 Distributions take their shape directly from the estimates, and they avoid the issues that arise from averaging.

Risk Model Integration

The next step in UC2 Analysis is to perform any final transformations needed to make the UC2 Distributions compatible with existing risk models and risk equations. Below are three examples that cover most existing use cases. The first two rely on weighted averages, and while that is not ideal, at least it is not an average of an average of an average, and so forth.

---

[6] https://www.flawofaverages.com

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

Binomial Inputs

Models that need binomial inputs can reduce a UC2 distribution to a single value using a weighted average. In this example, it reduces to 89%.

PERT Inputs

For models built to accept PERT estimates, provide the entire UC2 Distribution's range as the minimum and maximum values (1 - 100% in the example) and give the weighted average (89%) as the most-likely value.

Distribution Inputs

Models built to directly accept frequency distributions, such as PERT distributions, should be able to ingest a UC2 Distribution with little or no modification. Poisson and other distributions can also be mapped to US2 Distributions with little effort.

UC2 Risk Resolution

Most risk models do not have an explicit way to represent the state of Risk Resolution. While UC2 Analysis, described thus far, will improve models by correctly incorporating confidence and certainty, it gets lost when the UC2 Distribution is passed through its final transformation to become model inputs.

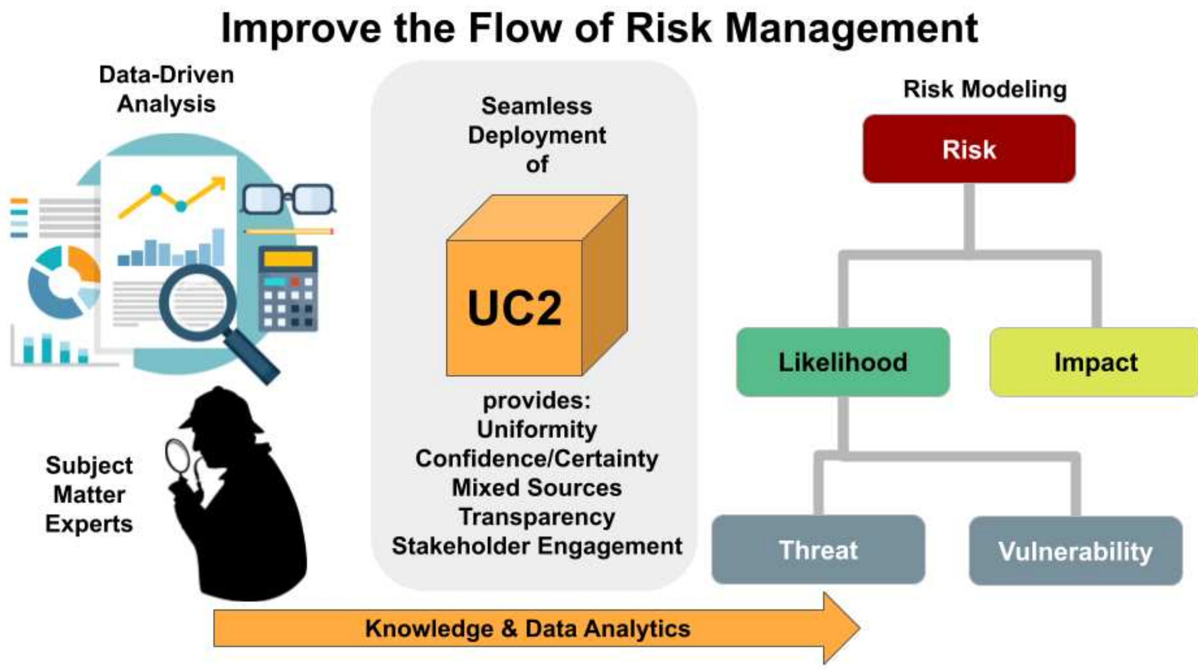
However, the UC2 Distributions themselves are useful to convey the state of Risk Resolution to stakeholders. With a bit of a layperson's explanation, they clearly illustrate the state of risk landscape "pixels" in terms of confidence and certainty through UC2 Distributions.

Conclusion

Without having to rip and replace existing risk management models or workflows, UC2 addresses several issues that plague current estimation techniques. Deployed between analysis and modeling, UC2 brings uniformity and interoperability that improve results and improve stakeholder engagement.

UC2 elevates risk management programs by increasing uniformity, transparency, and stakeholder engagement. Its unique features *accurately* capture confidence and certainty and improve interoperability between data-driven and expert-derived risk estimates and the models that use them.

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com



UC2, version one, is meant to seed a living and growing body of knowledge within risk management that unifies estimation techniques and improves risk analysis models. Critical feedback and opportunities to grow this body of knowledge include:

- Improved color-blind-safe palettes for UC2 visuals.
- Logo, icons and imagery.
- Improved labels and potential numeric definitions for Confidence Bands.
- Integration with SIPS, SLURPS[7], and Metalog[8].
- Blockchain databases for accountability, time-travel and shared ledgers[9].
- Tools to automate and simplify UC2 Field Usage and UC2 Analysis.
- Computational techniques to go beyond 10 segments in UC2 Scales.
- Adapt UC2 Scales to non-linear, exponential scales like earthquake magnitude.
- Application to maturity models like the NIST Cyber Security Framework (CSF), Cybersecurity Maturity Model Certification (CMMC), and so forth.
- The Common Vulnerability Scoring System (CVSS) might also benefit from integration with UC2, especially the aggregation of scores from multiple experts.
- UC2 may also be applicable to bodies of knowledge outside of risk management where variable precision and variable confidence estimates are used to approximate unknown values. Examples are healthcare, insurance, finance, etc.

[7] https://www.probabilitymanagement.org/
[8] http://metalogdistributions.com
[9] https://flur.ee

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

Academics and professionals are encouraged to contribute corrections, modifications and improvements directly to the author or through derivative works. It would also be helpful to know how UC2 is deployed in various contexts. While it is not a requirement, please let the author know if you choose to use UC2. Feedback is invaluable.

Released under the terms of the **Creative Commons Attribution-ShareAlike 4.0 International** license. You are *encouraged* to distribute, use, remix, or build upon this work. However, you *must* give credit — “Rob Arnold, Acorn Pass, LLC - https://AcornPass.com” — and you *must* share any derivatives under the same terms along with original credit.

Rob Arnold

https://AcornPass.com

Appendix A - ICD 203 Python Code

Run this in a python notebook to see how ICD 203 changes a uniform distribution.

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

# Define the labels and corresponding min/max values
labels = ['remote', 'highly improbable', 'improbable',
          'roughly even odds', 'probable', 'highly probable',
          'nearly certain']
#These are the bins defined by ICD 203
min_values = [1, 5, 20, 45, 55, 80, 95]
max_values = [5, 20, 45, 55, 80, 95, 99]

#Below are uniform bins, if you want to see what that does.
#min_values = [0, 14, 28, 42, 57, 71, 85]
#max_values = [14, 28, 42, 57, 71, 85, 99]

# Generate 10,000 random numbers between 1 and 99
np.random.seed(42)
random_numbers = np.random.randint(1, 99, size=10000)
# Comment out the line above and uncomment the next two lines to
create a normal distribution.
#random_numbers = np.random.normal(loc=49.5, scale=16.5,
size=10000).astype(int)
#random_numbers = np.clip(random_numbers, 0, 99)

# Plot the distribution of random numbers
plt.subplot(2, 1, 1)
plt.hist(random_numbers, bins=20, range=(1, 99),
edgecolor='black')
plt.xlabel('Number')
plt.ylabel('Count')
plt.title('Actual Estimation Distribution')

# Map random numbers to labels using the table
mapped_labels = []
for num in random_numbers:
    best_label = None
```

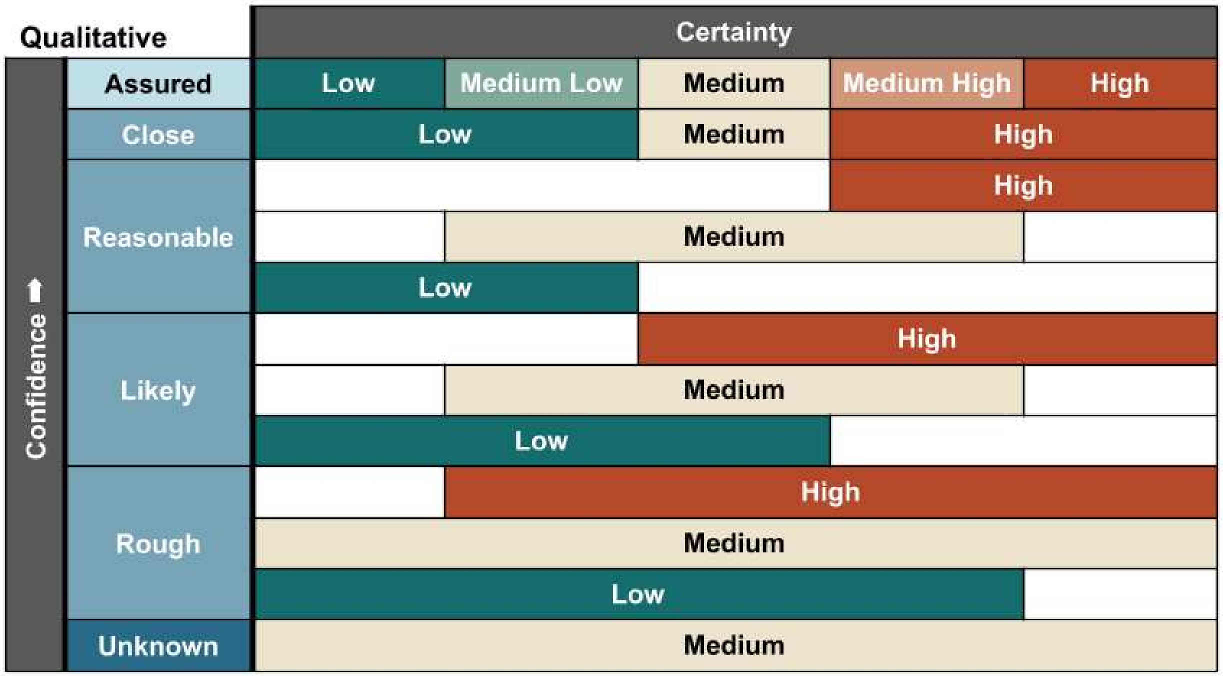

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

```
    for i, label in enumerate(labels):
        if min_values[i] <= num <= max_values[i]:
            best_label = (label, num)
    if best_label:
        mapped_labels.append(best_label)

# Create a dictionary to hold label counts in original order
label_counts = {label: 0 for label in labels}

# Fill up the counts in the dictionary
for label, _ in mapped_labels:
    label_counts[label] += 1

# Plot the distribution of labels
plt.subplot(2, 1, 2)
plt.bar(range(len(label_counts)), list(label_counts.values()))
plt.xlabel('Label')
plt.ylabel('Count')
plt.title('ICD 203 Distribution')
plt.xticks(range(len(label_counts)), list(label_counts.keys()),
rotation=45, ha='right')

plt.tight_layout()
plt.show()
```


CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

Appendix B - UC2 Qualitative Use Case

This example uses High, Medium, Low, etc. to illustrate the use of qualitative labels with a UC2 Scale.



Below is an example of data a risk analyst might receive from a facilitated expert elicitation session.

Session Date: 1/23/2023

Facilitator: J. Doe

Risk Scenario: How likely is a small retail business to encounter an incident of credit card theft in a given year?

Data Results:

Expert	Precision	Estimate	Facilitator Notes
R. Watson	Rough	High	
S. Jones	Reasonable	Medium	
Z. Darnell	Unknown	Medium	Scenario too volatile.
C. Hyland	Unknown	Medium	Expert lacks knowledge

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

K. Christian	Likely	High	
L. Gray	Close	High	

In the laboratory the risk analyst might consider dropping the C. Hyland estimate since the expert indicated they lack knowledge. This example it sets it to NULL.

This table, for compactness, uses UC2 numeric segment labels and is transformed as follows:

Expert	Estimates* (using numeric labels)	Facilitator Notes
R. Watson	2,3,4,5	
S. Jones	2,3,4	
Z. Darnell	1,2,3,4,5	Scenario too volatile.
C. Hyland	NULL	Expert lacks knowledge
K. Christian	3,4,5	
L. Gray	4,5	

** UC2 Scales define numeric labels in addition to word labels. Their use in this interim step illustrates how numeric labels can help represent results more compactly and improve integration with various analysis tools where numbers are easier to work with than word labels.*

Low (1)	***Medium Low (2)***	***Medium (3)***	***Medium High (4)***	***High (5)***

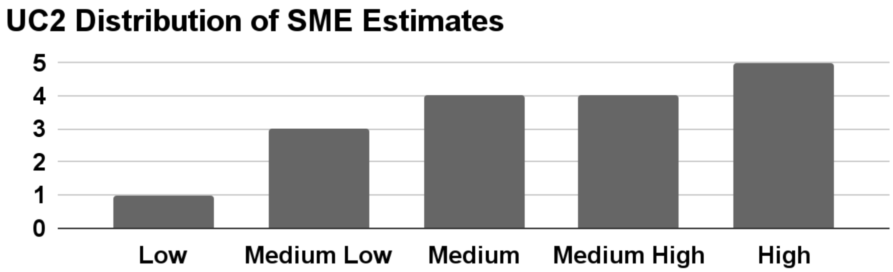
The data is then transformed by counting the number of times a segment of the full precision was part of an expert's estimate. Here the example switches back to using word labels for the estimates. UC2 was designed to allow this kind of switching back and forth to facilitate data manipulation (easier with numbers) and human readability (easier with words).

Estimate (Precise)	Count
High	1
Medium Low	3
Medium	4

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

Medium High	4
High	5

Finally, the results are graphed and interpreted as a distribution, which can be plugged into a risk model, Monte Carlo simulation, etc. Neither the expert nor the risk analyst needs to “fit” an abstract distribution to the data. The distribution simply emerges from the data, faithfully honoring confidence and certainty as it grows.



Note that dropping C. Hyland’s Unknown estimate did not change the shape of the distribution. If it were included, every bar in this graph would be one point higher, but the overall shape would not change. This illustrates how UC2 can cope with experts that realize in the field that they are not equipped to make a given estimate. In practice the risk analyst, perhaps with feedback from the facilitator, can choose to keep or discard Unknown estimates as warranted.

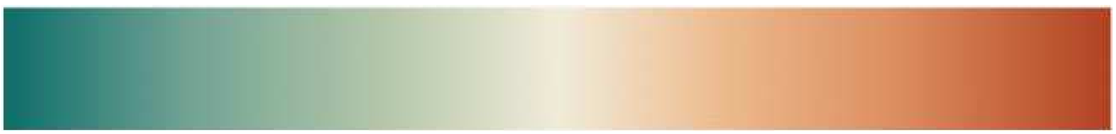
CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com

Appendix C - UC2 Color Pallets

Choosing an audience and context appropriate color pallet is a subject unto itself. A full discussion of the topic is out of scope for this paper. However, when designing the proposed solution, deliberate care was taken to select a pallet that is friendly to the roughly 5% of people that are colorblind. This adjustment aims to ensure colorblind experts feel confident using the tools and making good estimates.

The colors presented in this paper were derived from an article by Lisa Charlotte Muth[10]. The pallet was extracted from the gradients band below, which she presented in an elegant heatmap context.

Colorblind Gradient 3-Hues - Certainty (Midpoint Anchor)



Colorblind Gradient 1 Hue - Confidence

[10] https://blog.datawrapper.de/colors/ https://lisacharlottemuth.com/2016/04/22/Colors-for-DataVis/ https://blog.datawrapper.de/interpolation-for-color-scales-and-maps/

CC-BY-SA — Rob Arnold — **Acorn Pass, LLC** — https://AcornPass.com