Technomics

Better Decisions Faster

Uncertainty of Expert Judgment in Agile Software Sizing

ICEAA Conference
Pittsburgh, PA
Tuesday, May 17th, 2022
Mr. Peter J. Braxton, Technomics, Inc.

Fearful Asymmetry

Presented at the 2022 ICEAA Professional Development & Training Workshop: www.iceaaonline.com/pit2022

Abstract

Agile software estimating and planning often rely on expert judgment to assess the size of the development effort at various levels of granularity and stages of maturity. Previous research by the author quantified the inherent risk and uncertainty of the self-similar scales (e.g., T-shirt sizing) commonly used in these assessments. This paper expands those a priori mathematical results and empirically tests the accuracy of experts in applying those scales. It elucidates the ideal ratio to align with the desired confidence interval, and recommends feedback mechanisms to improve consistency.

Track: Management and Risk,

https://www.iceaaonline.com/pit22sessions/#MRTrack

Keywords: *Early Cost, Functional Requirements, Risk, Software, Uncertainty, Agile, Story Points*

Co-Authors

Peter Braxton

Subject Matter Expert

Performs cost and risk analysis for a number of federal clients. Has played integral roles in the development of both the SRDR and BCF 250 Applied Software Cost Estimating course at DAU. Dean of TTI Foundations of Cost Analysis curriculum. Current research interests include leveraging detailed Agile and DevOps data in forecasting program cost.

Dave Brown

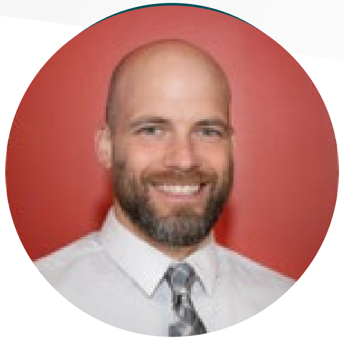
Subject Matter Expert

Performs cost estimating and analysis to DoD and DHS clients. Primary areas of expertise are IT and software estimating, with products such as life cycle cost analysis, applied cost estimating, independent cost assessment, cost research, program management support, modeling and simulation, data analysis, and database development.

Ken Rhodes

Senior Analyst

Performs cost analysis and acquisition decision support for DoD customers. Develops life-cycle cost estimates, cost / price assessments, and data visualization products for software and IT programs.



Alex Wekluk

Senior Analyst

Performs cost, risk, and technical analyses for the DoD and the IC. Secured a USPTO patent for Marine Corps weapon design work and earned the IC Meritorious Unit Citation for exemplary performance identifying cost-reduction measures.

???

Singer/Songwriter

Plays piano and sings. Member of Songwriters Hall of Fame (1992) and Rock and Roll Hall of Fame (1999), and Kennedy Center Honors recipient (2013). Wrote and recorded 33 Top 40 hits, including three #1's. Five Grammy awards, including Album of the Year.

Outline

- Problem Statement: Sizing Methods and T-Shirt Sizing Scales
- Thought Experiment: "Double or Half?"
- Problem Context: Self-Similar Scales in Agile Software Development
- Planning Poker, Fibonacci Numbers, and the Golden Mean
- Problem Context: Reliance on the Reluctant Expert
- Empirical Experiment: Analogized Scales

“Baby all the lights are turned on you/
Now you’re in the center of the stage”

“Everybody Loves You Now,” *Cold Spring Harbor*

The Basic Idea – Double or Half?!

- In the basic *Who Wants To Be a Millionaire* game, the dollar value (approximately) doubles for each question
  - $1,000 and $32,000 are “safe” plateaus
- Beyond $32,000, the contestant is faced with a choice:
  - Walk with the amount already earned, *or*
  - Go for the next question (“double”) *but*
  - Risk losing all but the $32K
- For the $64,000 Question – see what they did there?! – the losing side of the bet is precisely “half”

https://millionaire.fandom.com/wiki/Peter_Braxton

“We didn’t start the fire/
It was always burning, since the world’s been turning”

“We Didn’t Start the Fire,” *Storm Front*

Sizing Approaches – Definitions

- T-Shirt Sizing: Popularized by Agile Teams (S/M/L/XL)
- Planning Poker: Gamified technique to gather input from group
- Fibonacci Numbers: “borrowed from nature … allows relative sizing”
- Story Points: capture complexity, breadth, and risk
- Function Points (FP): based on logical data groups and processes
- Simple Function Points (SiFP): three transactional processes
- Source Lines of Code (SLOC): quantitative measurement

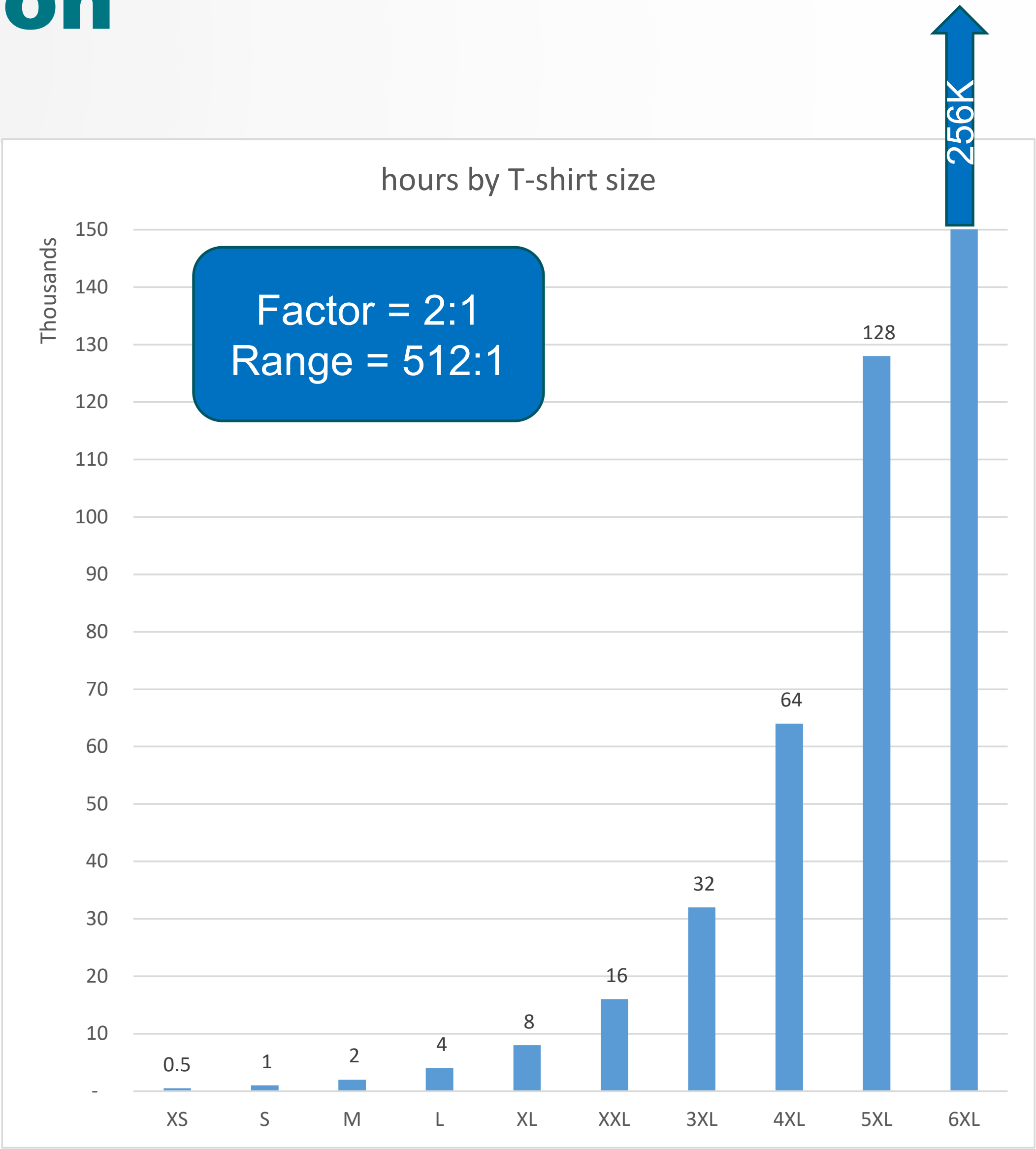
“an indication of effort”

“You may be right/ I may be crazy/
But it just might be a lunatic/ You’re looking for”

“You May Be Right,” *Glass Houses*

T-Shirt Sizing Risk – Introduction

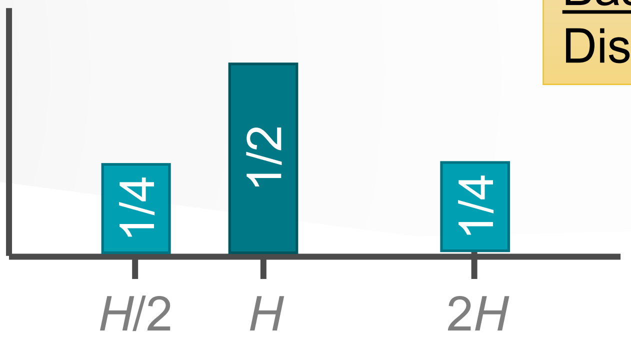
- T-Shirt Sizing is purposefully an exponential scale (aka logarithmic)
  - Similar to the use of Fibonacci numbers and "planning poker" in Agile
  - Other common logarithmic scales include Richter (earthquakes) and Decibel (sound)
- Going-in Risk position is that SME assessments could very easily be off by one T-shirt size in either direction
- Straightforward math leads to growth percentages and CVs under various distributional assumptions



T-Shirt Sizing Risk – General Framework

- Premise: A variation of the "double-or-half" thought experiment establishes a specific probability distribution
- Risk: Compute the **mean** of the probability distribution
  - Compare to the original point estimate (*H* hours) to establish a Cost Growth Factor (CGF), and equivalent **percent growth** (on average)
- Uncertainty: Compute the **variance** of the probability distribution
  - Compare standard deviation to the original point estimate ("pseudo CV") and estimate with growth to determine Coefficient of Variation (**CV**)
- Refinements:
  1. From discrete to *continuous* outcomes
  2. Incorporating degree of *confidence*
  3. Adjusting *beyond* "double-or-half" based on confidence
  4. Generalizing to ratios other than two

Naïve Uncertainty: Coin Flips



Base Case: Discrete

- Assume a Discrete distribution:
  - Most Likely = H hours, with a probability of 1/2
  - Max = $2H$ hours, with a probability of 1/4
  - Min = $H/2$ hours, with a probability of 1/4

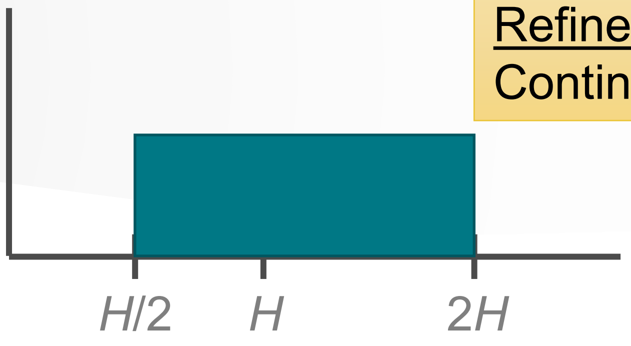
Coin flip #1: right or wrong
Coin flip #2: high or low

- Mean is expected value: $\sum_i x_i p_i = \left({}^1/_4\right)\left({}^H/_2\right) + \left({}^1/_2\right)(H) + \left({}^1/_4\right)(2H) = \frac{9H}{8} = \left(1 + \frac{1}{8}\right)H$
  - CGF = 1.125, or **12.5%** growth over point estimate
- Variance is expected value of square less square of expected value:

$$\sum_i x_i^2 p_i - \left[\sum_i x_i p_i\right]^2 = \left({}^1/_4\right)\left({}^{H^2}/_4\right) + \left({}^1/_2\right)(H^2) + \left({}^1/_4\right)(4H^2) - \left[\frac{9H}{8}\right]^2 = \frac{25H^2}{16} - \frac{81H^2}{64} = \left[\frac{\sqrt{19}}{8}H\right]^2$$

  - CV = **48.43%**

“Maximum” Uncertainty: Uniform



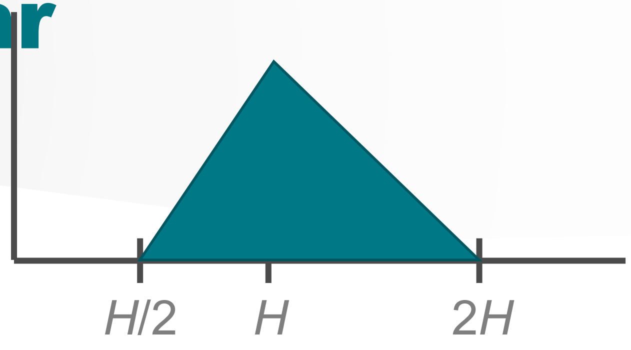
Refinement #1: Continuous

- Assume a Uniform distribution:
  - Max = 2*H* hours (next largest T-shirt size)
  - Min = *H*/2 hours (next smallest T-shirt size)
- Mean is average of Min/Max: $\frac{H/_2 + 2H}{2} = \frac{5H}{4} = \left(1 + \frac{1}{4}\right)H$
  - CGF = 1.25, or **25.0%** growth over point estimate
- Variance is range squared / 12:

$$\frac{\left(2H - H/_2\right)^2}{12} = \frac{9H^2}{4 \cdot 12} = \left[\sqrt{3} \cdot \frac{H}{4}\right]^2 = \left[\frac{\sqrt{3}}{4}H\right]^2$$

  - CV = **34.64%**

"Standard" Uncertainty: Triangular



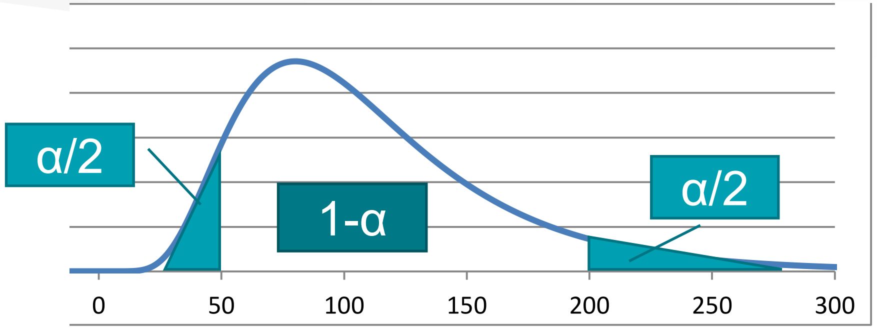
- Assume a Triangular distribution:
  - Most Likely = *H* hours (assessed T-shirt size)
  - Max = 2*H* hours (next largest T-shirt size)
  - Min = *H*/2 hours (next smallest T-shirt size)
- Mean is average of Min/ML/Max: $\frac{H/2 + H + 2H}{3} = \frac{7H}{6} = \left(1 + \frac{1}{6}\right)H$
  - CGF = 1.167, or **16.7%** growth over point estimate
- Variance is sum of squares less sum of pairwise products / 18:

$$\frac{(H/2)^2 + H^2 + (2H)^2 - H^2/2 - H^2 - 2H^2}{18} = \frac{7H^2/4}{18} = \frac{7H^2}{2 \cdot 36} = \left[\sqrt{\frac{7}{2}} \cdot \frac{H}{6}\right]^2 = \left[\frac{\sqrt{14}}{12} H\right]^2$$

  - CV = **26.73%**

"Standard" Risk: Lognormal

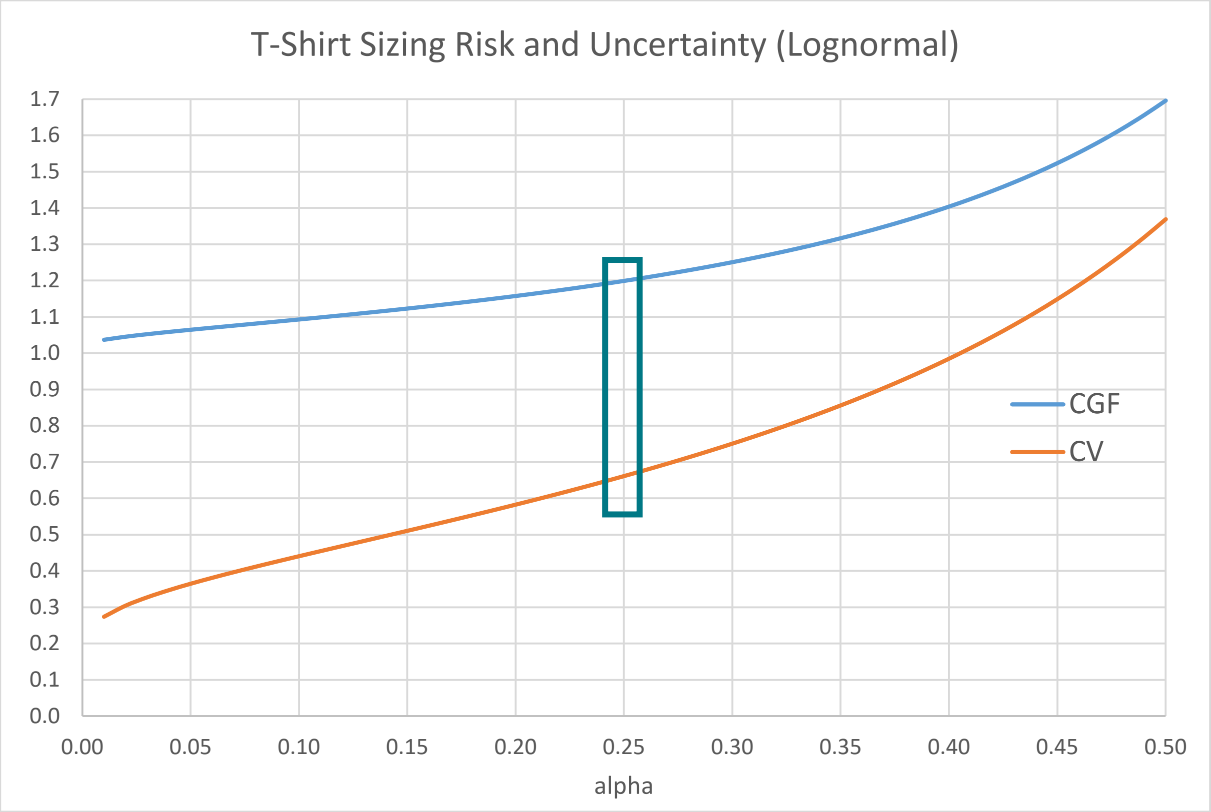
Refinement #3: Adjustment



- Assume a Lognormal distribution:
  - Median = *H* hours, with a probability of 1-α between *H*/2 and 2*H*
  - Right tail > 2*H* hours, with a probability of α/2
  - Left tail < *H*/2 hours, with a probability of α/2
- Confidence interval of related normal is: $(lnH - ln2, lnH, lnH + ln2)$
  - So that $\Phi^{-1}(1 - {}^{\alpha}/_{2}) = \frac{ln2}{\sigma}$ $\sigma = \frac{ln2}{\Phi^{-1}(1 - {}^{\alpha}/_{2})} = \frac{1}{log_2 e^{\Phi^{-1}(1 - {}^{\alpha}/_{2})}}$
- Mean of the lognormal is: $e^{\mu + \frac{\sigma^2}{2}}$
  - With a CGF of $e^{\frac{\sigma^2}{2}} = \sqrt{1 + CV^2}$ $CV = \sqrt{e^{\sigma^2} - 1}$

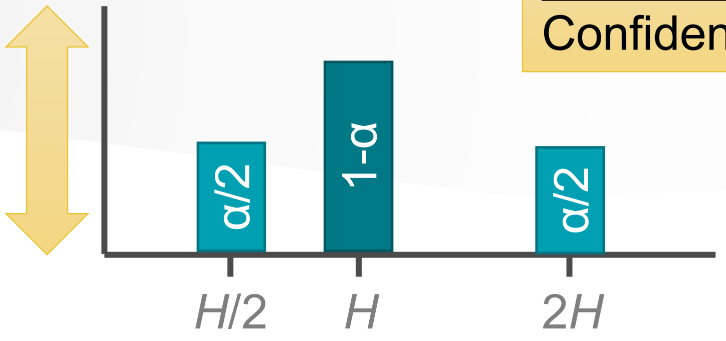
T-Shirt Sizing Risk – Lognormal (Illustrated)

- Graph illustrates increase in CGF and CV as percent chance outside the "double-or-half" range increases
  - Beyond α = 0.50 ("coin flip"), values increase rapidly



Generalization #1: Confidence

Refinement #2: Confidence



In previous example, α = 1/2

- Assume a Discrete distribution:
  - Most Likely = *H* hours, with a probability of 1-α
  - Max = 2*H* hours, with a probability of α/2
  - Min = *H*/2 hours, with a probability of α/2
- Mean is expected value: $\sum_i x_i p_i = (\alpha/2)(H/2) + (1-\alpha)(H) + (\alpha/2)(2H) = \left(1 + \frac{\alpha}{4}\right) H$
  - CGF = 1+(α/4), or **α/4** growth over point estimate
- Variance is expected value of square less square of expected value:

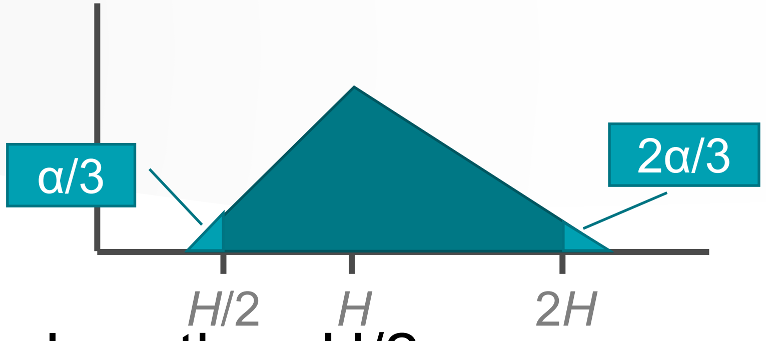
$$\sum_i {x_i}^2 p_i - \left[\sum_i x_i p_i\right]^2 = (\alpha/2)\left(H^2/4\right) + (1-\alpha)(H^2) + (\alpha/2)(4H^2) - \left[\left(1+\frac{\alpha}{4}\right)H\right]^2 =$$

$$\left(1+\frac{9\alpha}{8}\right)H^2 - \left(1+\frac{\alpha}{2}+\frac{\alpha^2}{16}\right)H^2 = \frac{10\alpha-\alpha^2}{16}H^2 = \left[\frac{\sqrt{10\alpha-\alpha^2}}{4}H\right]^2 \qquad CV = \frac{\sqrt{10\alpha-\alpha^2}}{4+\alpha}$$

T-Shirt Sizing Risk – Discrete (Illustrated)

- Graph illustrates range between always right (α=0) and always wrong (α=1), with a coin flip to determine low or high
  - Max growth is 25%
  - Max CV is 60%

Triangular Expanded – Proportional

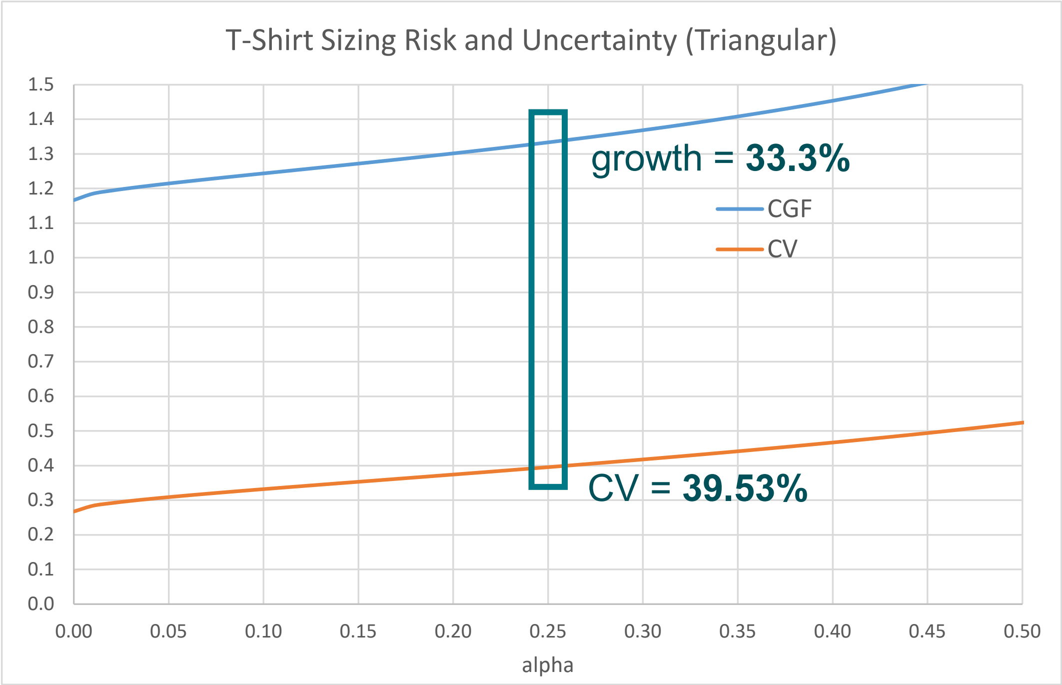


- Assume that the interval (H/2, 2H) encapsulates only (1- α) of the probability
  - That is, there is probability α of being greater than 2H or less than H/2
  - This can be split proportionally or equally
- Proportional puts $\frac{2\alpha}{3}$ above and $\frac{\alpha}{3}$ below

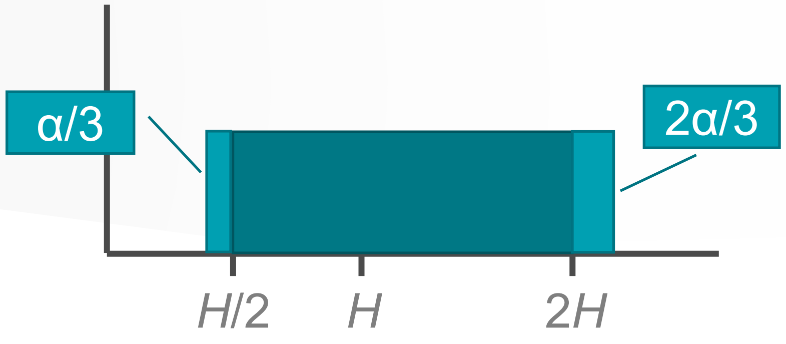
$$\mu = \left[\left(1-\frac{\sqrt{\alpha}}{1-\sqrt{\alpha}}\right)\frac{H}{2}+H+\left(2+\frac{\sqrt{\alpha}}{1-\sqrt{\alpha}}\right)H\right]\Big/3 = \left(1+\frac{1}{6-6\sqrt{\alpha}}\right)H$$

- Variance:

$$\left[\frac{\sqrt{\frac{7-4\sqrt{\alpha}}{2}}}{6-6\sqrt{\alpha}}H\right]^2 \qquad CV = \frac{\sqrt{\frac{7-4\sqrt{\alpha}}{2}}}{7-6\sqrt{\alpha}}$$



Proportional Tails – Uniform

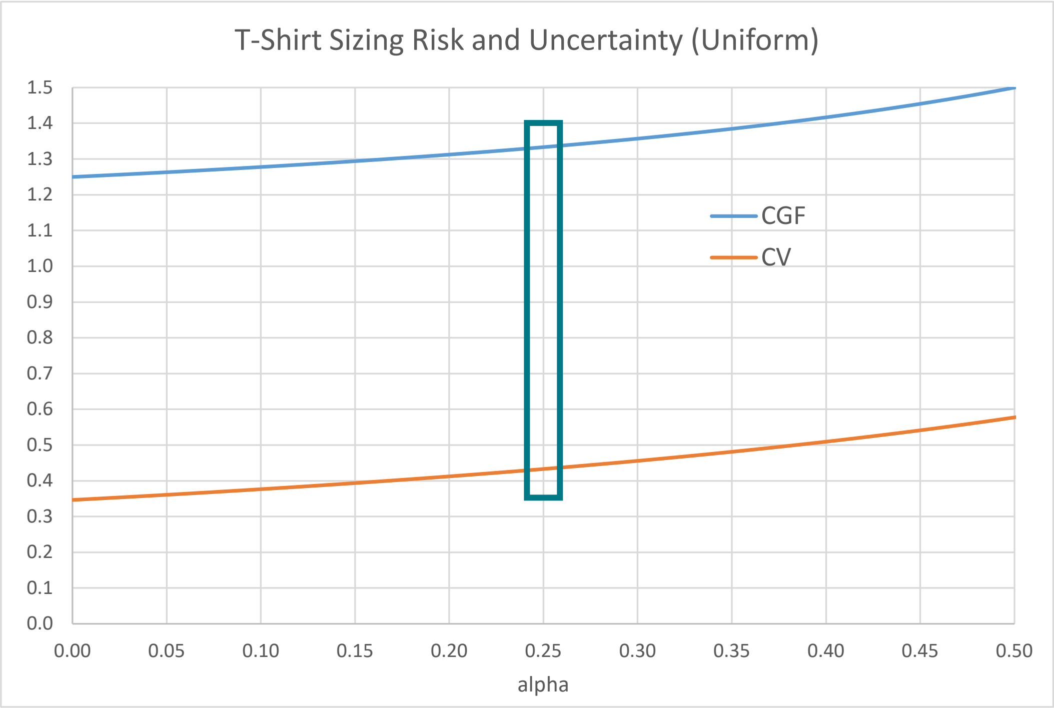


- Assume that the interval (H/2, 2H) encapsulates only (1-α) of the probability
  - That is, there is probability α of being greater than 2H or less than H/2
  - This can be split proportionally or equally
- Proportional puts $\frac{2\alpha}{3}$ above and $\frac{\alpha}{3}$ below

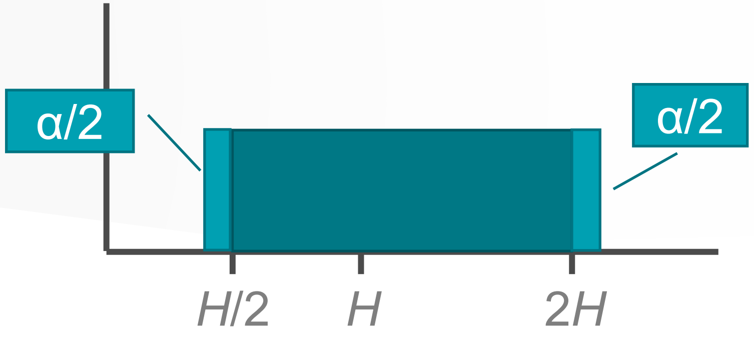
$$\mu = \left[\frac{(1-2\alpha)}{(1-\alpha)}\frac{H}{2} + \frac{(2-\alpha)}{(1-\alpha)}H\right]\Big/_2 = \frac{5-4\alpha}{4-4\alpha}H = \left(1+\frac{1}{4-4\alpha}\right)H$$

- Variance is range squared / 12:

$$\frac{(3H)^2}{12[2(1-\alpha)]^2} = \left[\frac{\sqrt{3}}{4-4\alpha}H\right]^2$$



Symmetric Tails – Uniform

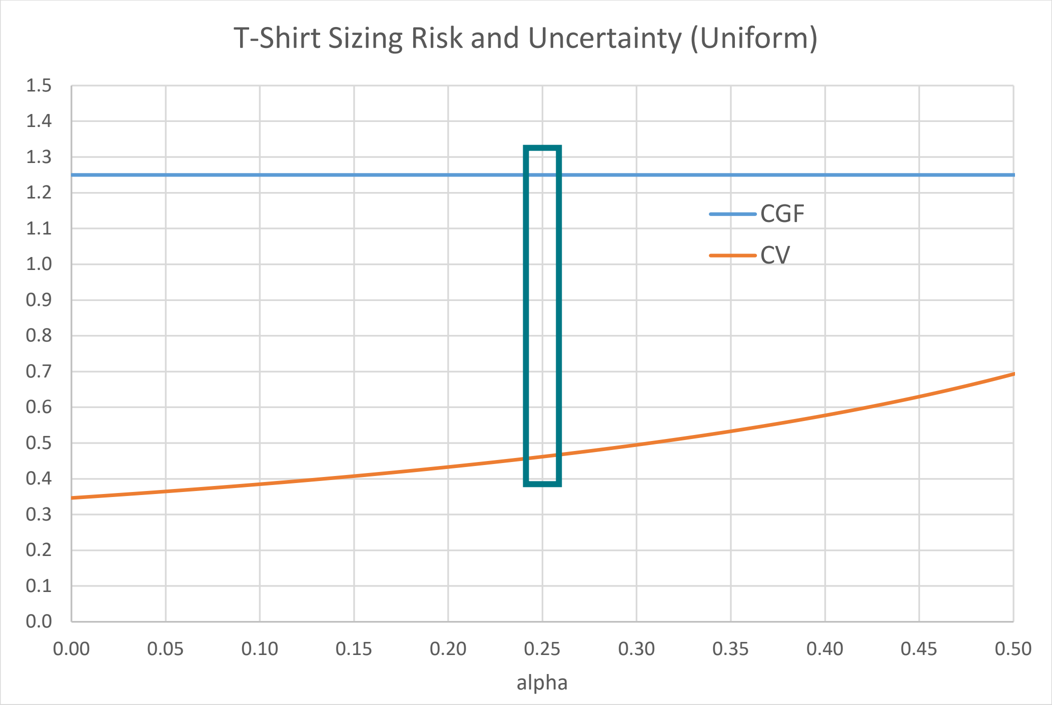


- Assume that the interval (H/2, 2H) encapsulates only (1-α) of the probability
  - That is, there is probability α of being greater than 2H or less than H/2
  - This can be split proportionally or equally
- Equal puts $\frac{\alpha}{2}$ above and $\frac{\alpha}{2}$ below

$$\mu = \left[\frac{(2-5\alpha)}{(4-4\alpha)}H + \frac{(8-5\alpha)}{(4-4\alpha)}H\right]\Big/_2 = \frac{5}{4}H = \left(1+\frac{1}{4}\right)H$$

- Variance is range squared / 12:

$$\frac{(6H)^2}{12[4(1-\alpha)]^2} = \left[\frac{\sqrt{3}}{4-4\alpha}H\right]^2$$



Risk and Uncertainty by Confidence

- For confidence (1-α), we can express CGF and CV as a function of α
  - Generally, we would assume α < 0.50 (i.e., no worse than coin flip)

	Growth %	CV	Growth % (α = 0.25)	CV (α = 0.25)
Discrete (Generalized)	$\frac{\alpha}{4}$	$\frac{\sqrt{10\alpha - \alpha^2}}{4+\alpha}$	6.2%	36.74%
Lognormal	$\sqrt{1+CV^2} - 1$	$\sqrt{e^{\sigma^2} - 1}$	19.9%	66.16%
Uniform (Proportional)	$\frac{1}{4-4\alpha}$	$\frac{\sqrt{3}}{5-4\alpha}$	33.3%	43.30%
Uniform (Equal)	$\frac{1}{4}$	$\frac{\sqrt{3}}{5-5\alpha}$	25.0%	46.19%
Triangular (Proportional)	$\frac{1}{6-6\sqrt{\alpha}}$	$\frac{\sqrt{\frac{7-4\sqrt{\alpha}}{2}}}{7-6\sqrt{\alpha}}$	33.3%	39.53%

$$\sigma = \frac{1}{log_2 e^{\Phi^{-1}(1-\alpha/2)}}$$

Planning Poker and Fibonacci Numbers

- Alternate sizing method is Planning Poker
  - Commonly uses Fibonacci numbers for sizing via Story Points
  - In some alternative formulations, larger sizes are replaced with “rounder” numbers
  - Often visualized using fruits!
- Combines “additive” and “multiplicative” features:
  - Sum of any two consecutive sizes is equal to the next largest size
  - Ratio of consecutive sizes *approaches* a constant
- Fibonacci numbers are the sequence starting with 1 and 1, and whose subsequent entries are the sum of the two previous numbers
  - 2 = 1+1, 3 = 1+2, 5 = 2+3, 8 = 3+5, 13 = 5+8, 21 = 8+13, 34 = 13+21, etc.

Fibonacci Numbers and the Golden Ratio

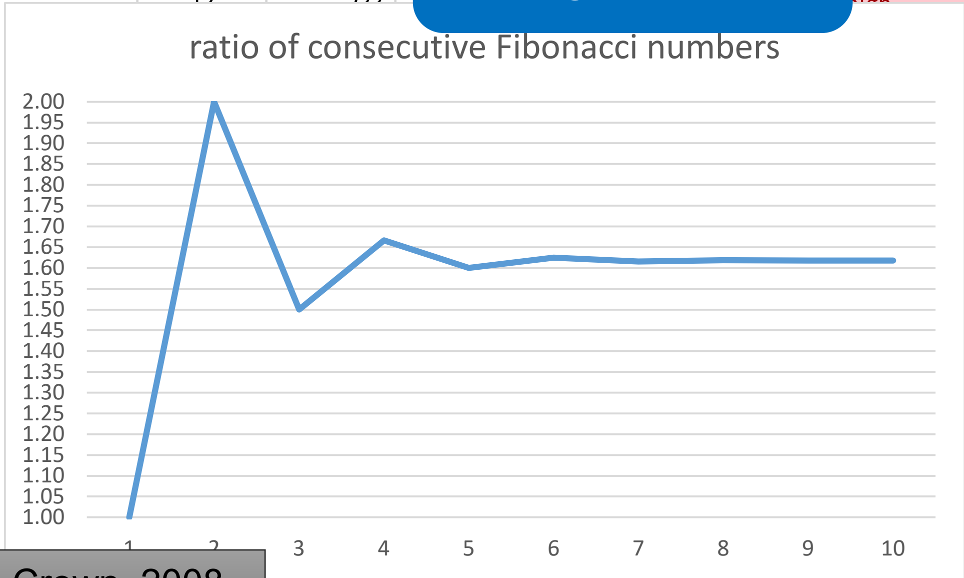
- Because the Fibonacci sequence is additive, the ratio between consecutive terms is *not* constant
- However, the ratio does quickly converge to a constant
  - It turns out that this is the Golden Ratio!

$$\phi = \frac{1+\sqrt{5}}{2} = 1.618\ldots$$

$$F_n = \frac{1}{\sqrt{5}}\left[\phi^n - (1-\phi)^n\right]$$

n	Fn	closed form	ratio	low/high
1	1	1		
2	1	1	1.000000	low
3	2	2	2.000000	high
4	3	3	1.500000	low
5	5	5	1.666667	high
6	8	8	1.600000	low
7	13	13	1.625000	high
8	21	21	1.615385	low
9	34			high
10	55			low
11	89			high
12	144			low
	832,040	832,040	1.618034	low

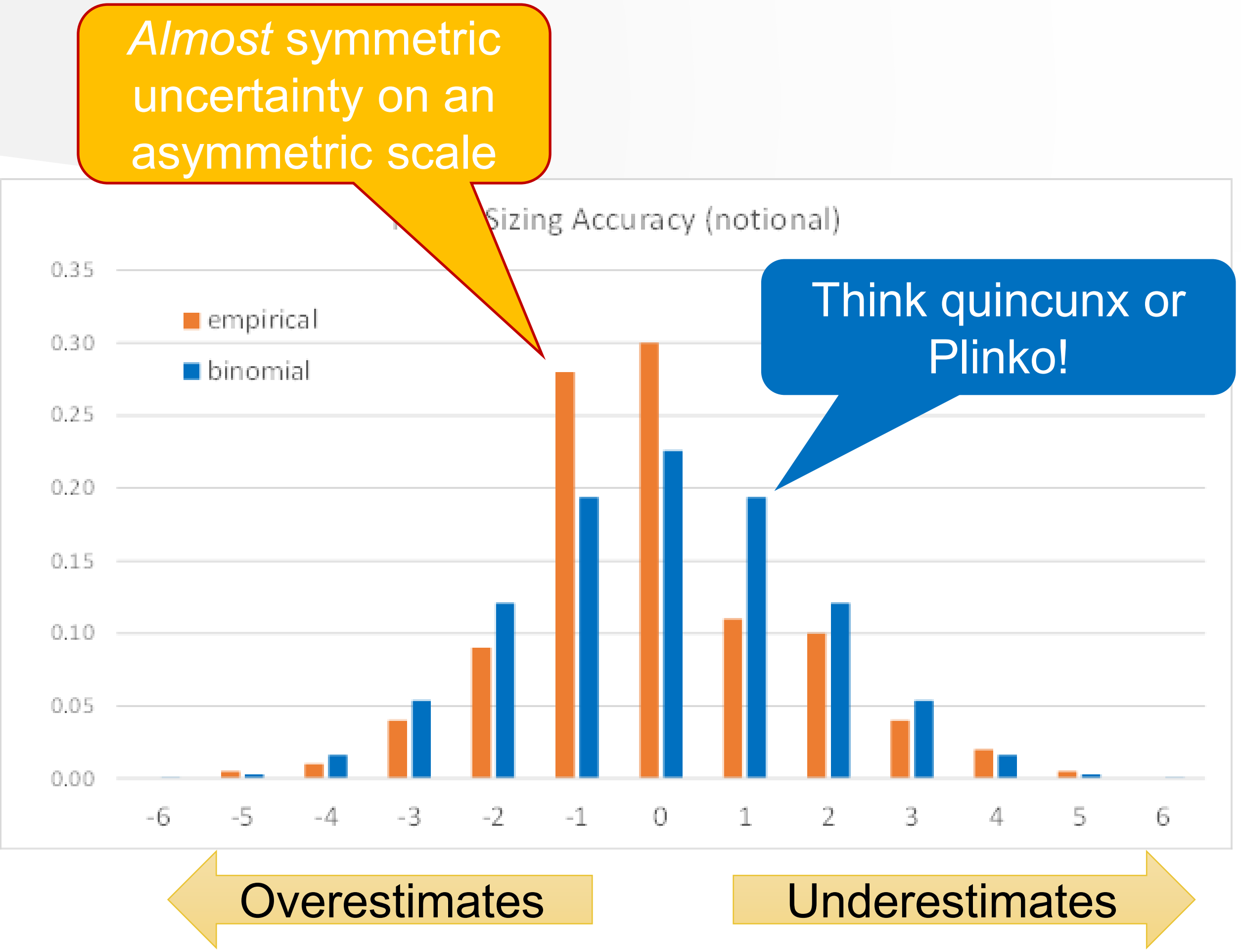
Factor = 1.618:1
Range = 144:1



The Golden Ratio: The Story of PHI, the World's Most Astonishing Number, Mario Livio, Crown, 2008.

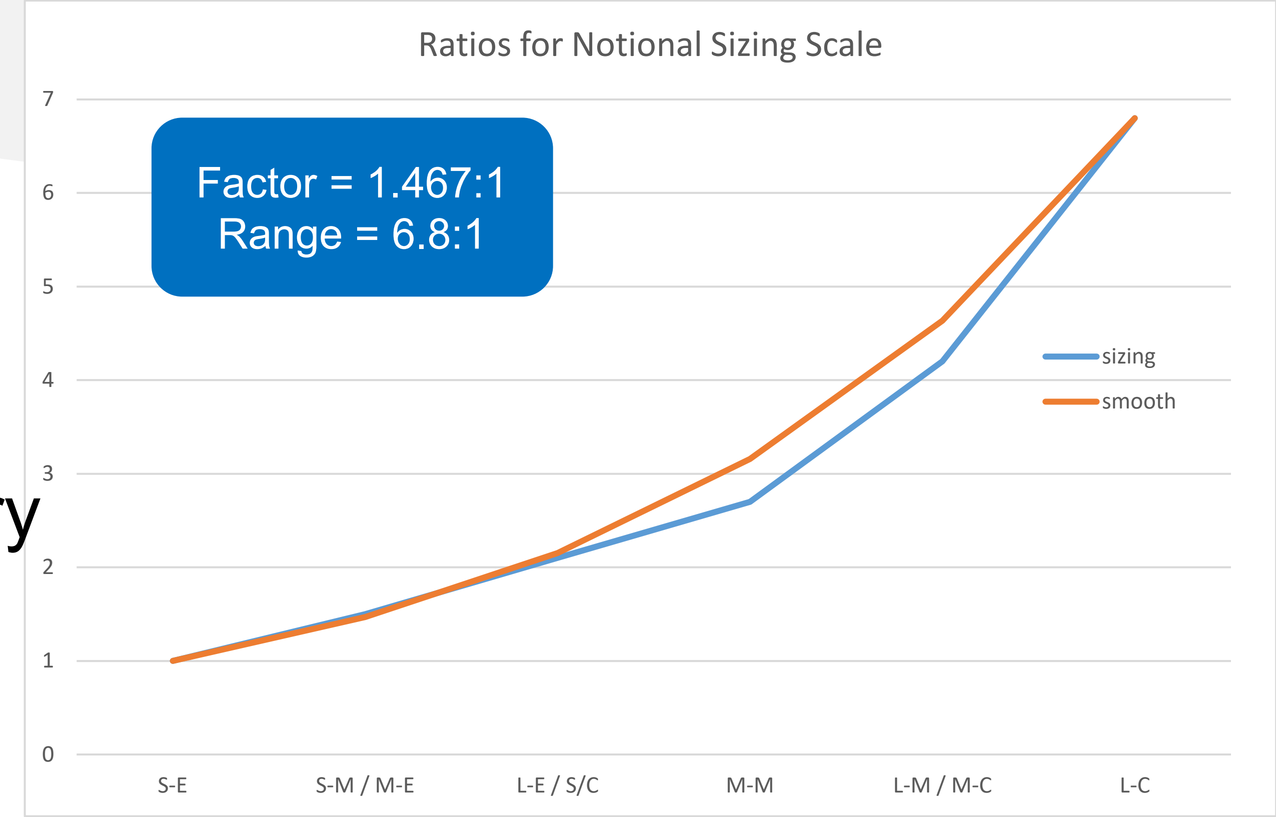
Micro-Sizing Accuracy

- As presented, T-shirt sizing is Macro level, whereas Fibonacci numbers are Micro level
- Still gathering empirical evidence on Macro-sizing accuracy
  - Initial evidence for Micro-sizing is largely consistent with hypothesized model
  - Except there may be *many* coin flips, not just one…



Notional Sizing Model

- Incorporates Size and Complexity
  - Small, Medium, Large
  - Easy, Moderate, Complex
- Additional assumption of symmetry maps 3 x 3 model to 6-point scale
  - Total range 1 : 6.8
  - Average "notch" ratio 1.467



Sked (mo)	S	M	L
E	12	15	18
M	15	18	21
C	18	21	24

LOE (FTE)	S	M	L
E	2.5	3	3.5
M	3	4.5	6
C	3.5	6	8.5

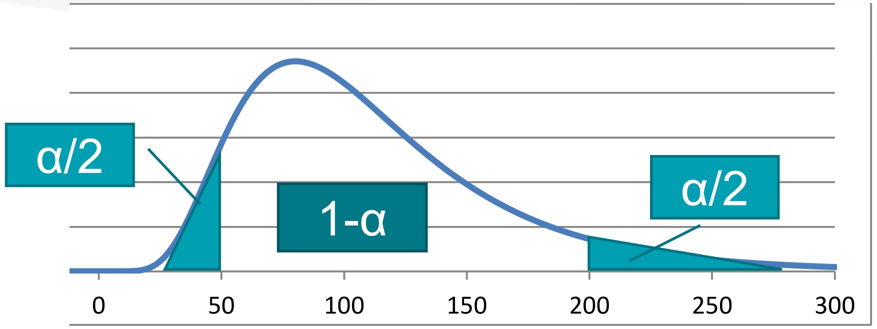
effort (PM)	S	M	L
E	30	45	63
M	45	81	126
C	63	126	204

effort (relative)	S	M	L
E	37.0%	55.6%	77.8%
M	55.6%	100.0%	155.6%
C	77.8%	155.6%	251.9%

Generalized Sizing Risk – Lognormal

Refinement #4: Generalized Ratio

- Assume a Lognormal distribution:
  - Median = *H* hours, with a probability of 1-α between *H*/*R* and *RH*
  - Right tail > *RH* hours, with a probability of α/2
  - Left tail < *H*/*R* hours, with a probability of α/2

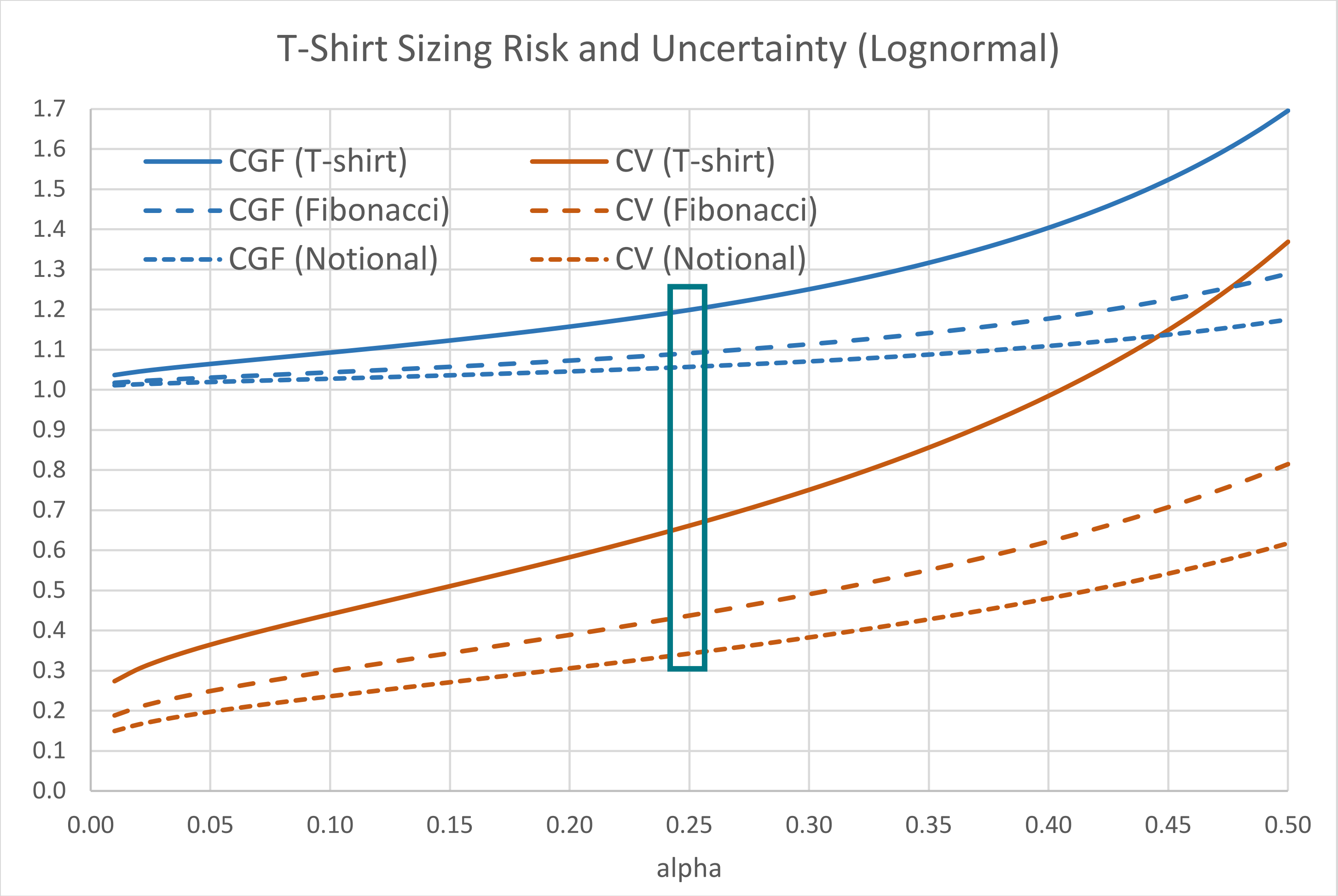


- Confidence interval of related normal is: $(lnH - lnR, lnH, lnH + lnR)$
  - So that $\Phi^{-1}(1 - {}^{\alpha}/_{2}) = \frac{lnR}{\sigma}$ $\quad \sigma = \frac{lnR}{\Phi^{-1}(1 - {}^{\alpha}/_{2})} = \frac{1}{log_R e^{\Phi^{-1}(1 - {}^{\alpha}/_{2})}}$

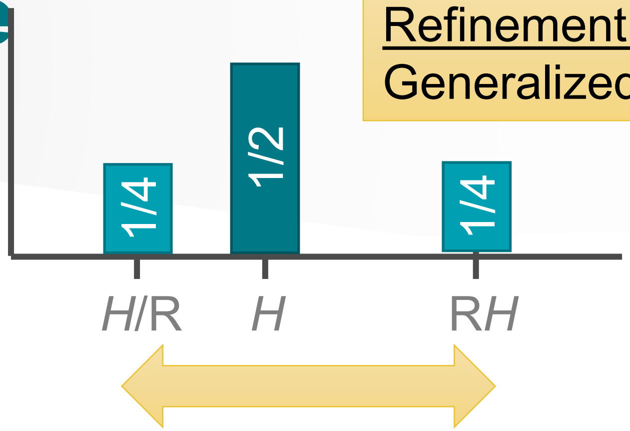
- Mean of the lognormal is: $e^{\mu + \frac{\sigma^2}{2}}$
  - With a CGF of $e^{\frac{\sigma^2}{2}} = \sqrt{1 + CV^2}$ $\quad CV = \sqrt{e^{\sigma^2} - 1}$

Generalized Risk – Lognormal (Illustrated)

- Common factors shown for T-shirt sizing (2.000), Fibonacci (1.618), and Notional (1.467)



Generalized Sizing Risk – Discrete



Refinement #4: Generalized Ratio

- Assume a Discrete distribution:
  - Most Likely = H hours, with a probability of 1/2
  - Max = RH hours, with a probability of 1/4
  - Min = H/R hours, with a probability of 1/4
- Mean is expected value:

$$\sum_i x_i p_i = \left(1/4\right)\left(H/R\right) + \left(1/2\right)(H) + \left(1/4\right)(RH) = \frac{1}{R}\left(\frac{R+1}{2}\right)^2 H = \left[1 + \frac{1}{R}\left(\frac{R-1}{2}\right)^2\right] H$$

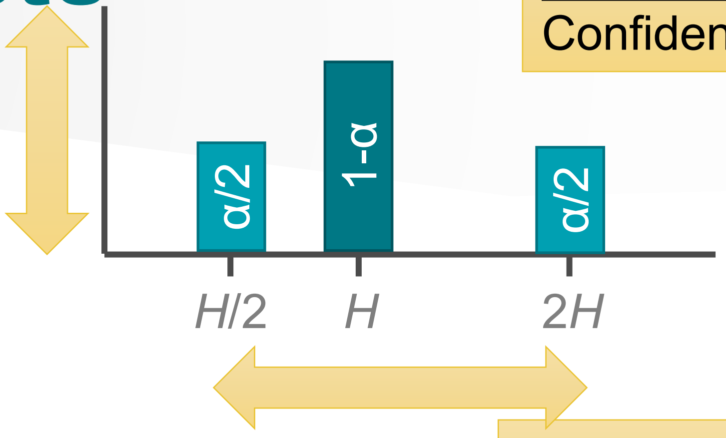
- Variance is expected value of square less square of expected value:

$$\sum_i {x_i}^2 p_i - \left[\sum_i x_i p_i\right]^2 = \frac{1}{4}\left(\frac{H}{R}\right)^2 + \frac{1}{2}H^2 + \frac{1}{4}(HR)^2 - \frac{1}{R^2}\left(\frac{R+1}{2}\right)^4 H^2 =$$

$$\left[\frac{3R^4 - 4R^3 + 2R^2 - 4R + 3}{(4R)^2}\right] H^2 = \left[\frac{R-1}{4R}\sqrt{3R^2 + 2R + 3}\right]^2 H^2$$

$$CV = \frac{R-1}{(R+1)^2}\sqrt{3R^2 + 2R + 3}$$

Generalized Sizing Risk – Discrete



Refinement #2: Confidence

Refinement #4: Generalized Ratio

- Assume a Discrete distribution:
  - Most Likely = H hours, with a probability of 1-α
  - Max = RH hours, with a probability of α/2
  - Min = H/R hours, with a probability of α/2
- Mean is expected value:

$$\sum_i x_i p_i = \left({}^{\alpha}/_{2}\right)\left({}^{H}/_{R}\right) + (1-\alpha)H + \left({}^{\alpha}/_{2}\right)(RH) = \frac{\alpha + 2(1-\alpha)R + \alpha R^2}{2R} H = \left[1 + \alpha\frac{(R-1)^2}{2R}\right] H$$

- Variance is expected value of square less square of expected value:

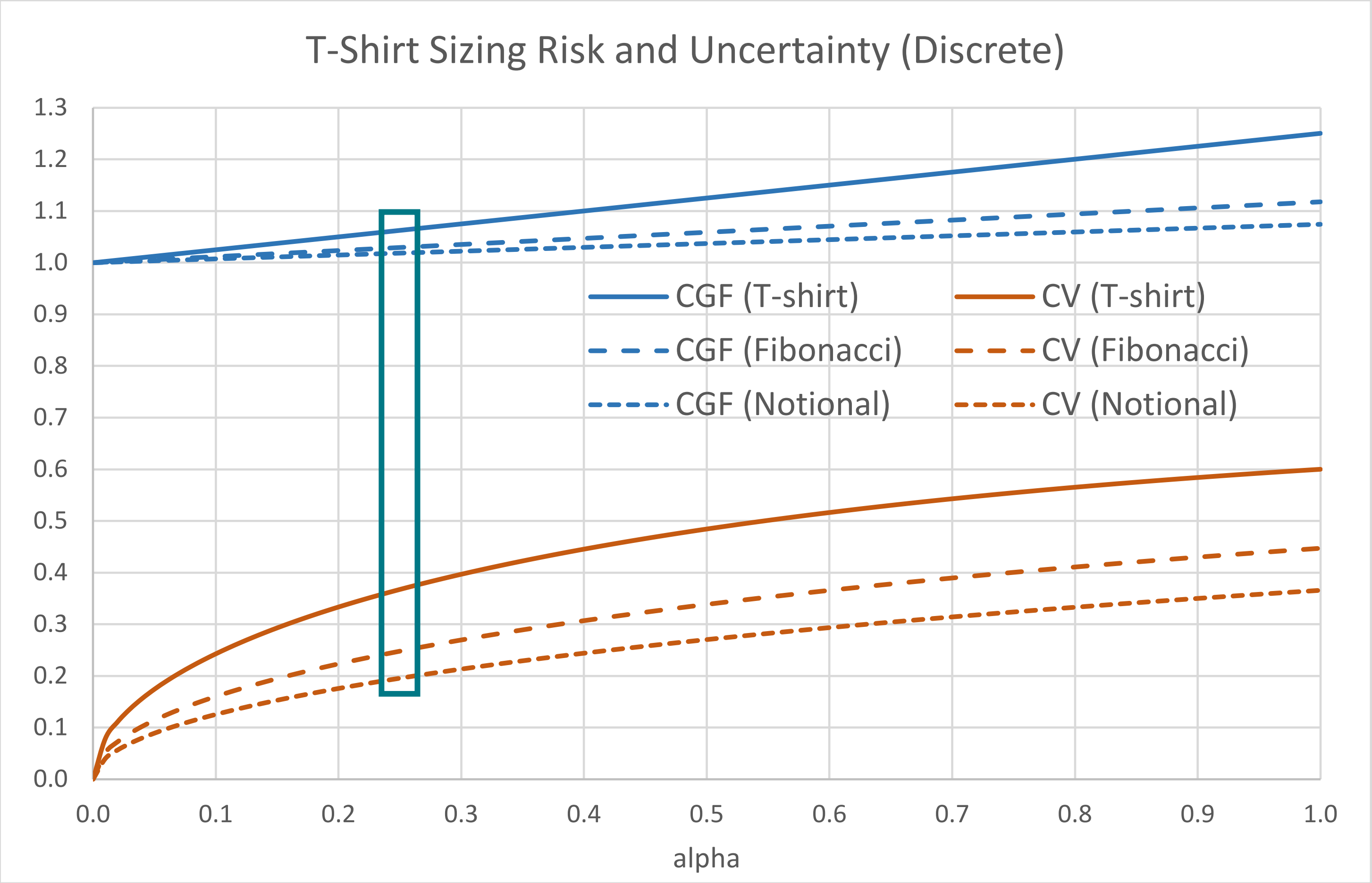
$$\sum_i {x_i}^2 p_i - \left[\sum_i x_i p_i\right]^2 = \frac{\alpha}{2}\left(\frac{H}{R}\right)^2 + (1-\alpha)H^2 + \frac{\alpha}{2}(HR)^2 - \left(\frac{\alpha - 2(1-\alpha)R + \alpha R^2}{2R}\right)^2 H^2 =$$

$$\left[\left(\frac{R-1}{2R}\right)\sqrt{\alpha[(2-\alpha)R^2 + 2\alpha R + (2-\alpha)]}\right]^2 H^2$$

$$CV = \frac{R-1}{2R + \alpha(R-1)^2}\sqrt{\alpha[(2-\alpha)R^2 + 2\alpha R + (2-\alpha)]}$$

Generalized Risk – Discrete (Illustrated)

- Common factors shown for T-shirt sizing (2.000), Fibonacci (1.618), and Notional (1.467)



“Honesty is such a lonely word/ Everyone is so untrue
Honesty is hardly ever heard/ But mostly what I need from you”

“Honesty,” *52nd Street*

Self-Similar Scales and the Ideal Ratio

- Self-similar scales are fractal in that misestimation will result in growth (or reduction) by the same ratio regardless of position on the scale
- Candidate ratios (R):
  - Two (2.0) – T-shirt Sizing
  - Phi (1.618…) – Planning Poker (Fibonacci numbers)
  - e (2.718…) – base of the exponential function that is its own derivative!
- It is proposed that these approximately bound the reasonable set of choices
- Related issue is “top-down” vs. “bottom-up”
  - Size more complex pieces of work as whole (initially) or force decomposition

Empirical Testing of Scales

- Approach used in previous paper on use of SME's in Cost and Risk
  - Both knowable but unknown past events (e.g., box office gross of *Avengers: Endgame*) and unknown future events (e.g., box office gross of *Thor: Love and Thunder*)
- Instead of asking for three-point estimates, ask for single best guess (closest value) from self-similar scale
  - Does gradation of scale affect accuracy of assessments?
- Expertise in subject area vs. expertise in uncertainty assessments

"Teaching Pigs to Sing: Improving Fidelity of Assessments from Subject Matter Experts (SMEs)," Peter Braxton and Richard Coleman, ICEAA Washington Chapter, June, 2012.

Expert Judgment vs. Expert Opinion

- Expert Opinion = estimate is presented as a direct assessment by SME with no apparent basis
- Expert Judgment = SME uses or interprets data as the basis of the estimate, or at worst makes a direct assessment as to the scope on which the estimate is based (e.g., software sizing!)
- It is hypothesized that sizing and similar assessments can be improved by labeling each notch on the scale with an actual example reflecting that approximate size
  - Transcends Expert Opinion with a sort of a "stealth" Analogy
  - Heights of mountains, e.g., could be used in empirical assessment

Cost Estimating Body of Knowledge (CEBoK®), Module 2 "Cost Estimating Techniques," ICEAA, 2013.

From Single-Point Analogy to Analogized Scales

- Benefits of an explicit Basis and Rationale:
  - *Independently verified* before the fact
  - *Empirically measured* after the fact
- "Analogizing" the self-similar scale
  - Augment or replace numerical values with historical examples
  - Similar to Mohs scale (mineral hardness), Beaufort scale (wind)
- Double "stealth"
  - Analogy estimate masquerading as Expert Opinion/Judgment
  - Three-point estimate masquerading as one-point estimate

Experimental Formulation

- Six basic treatments (proposed)
  - Scale labeling: numbers only, analogies only, or both
  - Scale ratio: 1.5 or 2.0
- Experiment #1: Heights of Mountains
  - Unknown but knowable, generally relatable

scale (ft)	mountain	location	elevation (ft)
500	Driskill Mountain	Louisiana	535
1,000	Woodall Mountain	Mississippi	807
2,000	Mount Arvon	Michigan	1,979
4,000	Black Mountain	Kentucky	4,145
8,000	Guadelupe Peak	Texas	8,751
16,000	Mont Blanc	France	15,774
32,000	Mount Everest	Nepal	29,031

scale (ft)	mountain	location	elevation (ft)
1,000	Woodall Mountain	Mississippi	807
1,500	Crown Mountain	St. Thomas, USVI	1,555
2,250	Eagle Mountain	Minnesota	2,302
3,375	Mount Davis	Pennsylvania	3,213
5,063	Black Mesa	Oklahoma	4,975
7,594	Black Elk Peak	South Dakota	7,244
11,391	Mount Hood	Oregon	11,249
17,086	Pico Pan de Azucar	Colombia	17,060
25,629	Nanda Devi	India	25,643

Additional Experiments

- Experiment #2: Box Office Gross of Films
  - Popular films from 1990-2019 (pre-pandemic) per Box Office Mojo
  - *Not* inflation-adjusted
  - Representative of macro-level sizing
  - For a $1M to $1B range, 11-point scale (R = 2.0) or 17-point scale (R = 1.5)
- Experiment #3: Driving Distances
  - From Technomics HQ in Arlington, VA, to local and interstate destinations
  - Test the fractal nature of risk

Conclusion

- More remains to be explored on empirical testing
- The bottom line is that significant risk and uncertainty are inherent in these self-similar sizing scales *even if we are off by no more than one size in either direction*

	Confidence	Growth %	CV
Discrete	$\alpha = 0.50$	12.5%	48.43%
Uniform	$\alpha = 0.00$	25.0%	34.64%
Triangular	$\alpha = 0.00$	16.7%	26.73%
Discrete	$\alpha = 0.25$	6.2%	36.74%
Lognormal	$\alpha = 0.25$	19.9%	66.16%
Uniform (Proportional)	$\alpha = 0.25$	33.3%	43.30%
Uniform (Equal)	$\alpha = 0.25$	25.0%	46.19%
Triangular (Proportional)	$\alpha = 0.25$	33.3%	39.53%

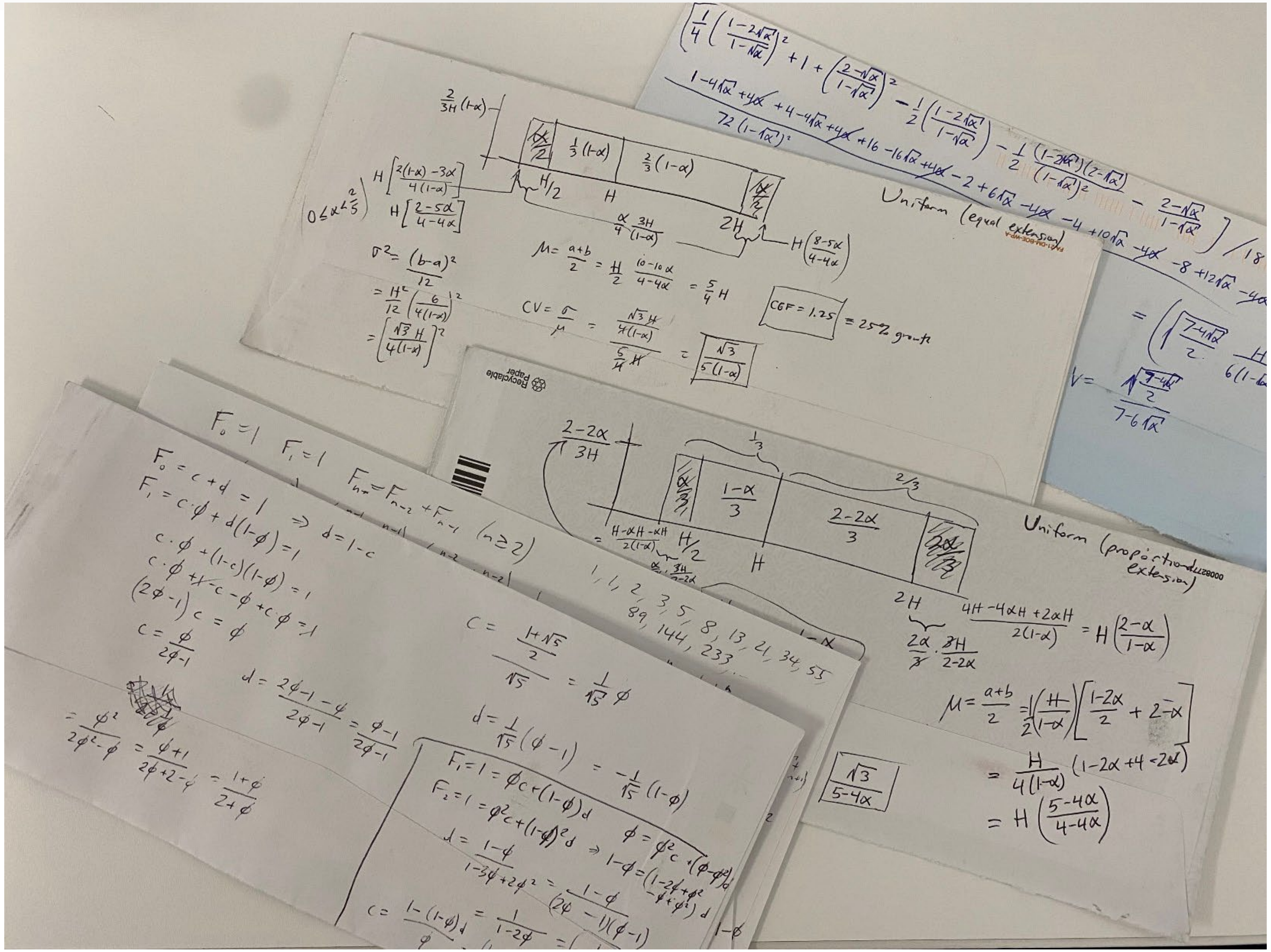
Bibliography

- Braxton, P. J. (2021). Inherent Risk and Uncertainty of Self-Similar Sizing Scales in Agile Software Development, or "Does This T-Shirt Make My Estimate Look Big?". *IT-CAST Proceedings.* Washington, DC: DHS.
- Braxton, P. J., & Coleman, R. L. (2012). Teaching Pigs to Sing: Improving Fidelity of Assessments from Subject Matter Experts (SMEs). *ICEAA Chapter Luncheon Proceedings.* Washington, DC: ICEAA Washington Chapter.
- Braxton, P. J., & Sayer, L. H. (2013). Probability Distributions for Risk Analysis. *ICEAA Conference Proceedings.* New Orleans, LA: ICEAA.
- *Cost Estimating Body of Knowledge (CEBoK).* (2013). Annandale, VA: ICEAA.
- Gellatly, W., Braxton, P. J., Brown, D. H., Jones, L. F., & Wekluk, R. A. (2022). Dynamic Software Effort Estimation: How SWEET It Is! *ICEAA Conference Proceedings.* Pittsburgh, PA: ICEAA.
- Kahneman, D. (2011). *Thinking, Fast and Slow.* New York, NY: Farrar, Straus and Giroux.
- Kosmakos, C., & Brown, D. H. (2022). Are We Agile Enough to Estimate Agile Software Development Costs? *ICEAA Conference Proceedings.* Pittsburgh, PA: ICEAA.
- Livio, M. (2008). *The Golden Ratio: The Story of PHI, the World's Most Astonishing Number.* Crown.
- Radigan, D. (2022, February 24). *Story points and estimation*. Retrieved from Atlassian Agile Coach: https://www.atlassian.com/agile/project-management/estimation

“It’s a pretty good crowd for Saturday
And the manager gives me a smile
‘Cause he knows that it’s me they’ve been comin’ to see
To forget about life for a while”

“Piano Man,” *Piano Man*

Coda – The Proverbial Cocktail Napkin(s)



Uncertainty of Expert Judgment in Agile Software Sizing

Back-Up

Fibonacci Numbers Closed-Form Formula

- A closed-form formula can be derived, which will easily demonstrate the convergence property
- Suppose a relationship of the form

$$F_n = c \cdot a^n + d \cdot b^n$$

- Then the recursive formula will be satisfied if *a* and *b* are roots of the quadratic

$$F_n + F_{n+1} = c \cdot a^n + d \cdot b^n + c \cdot a^{n+1} + d \cdot b^{n+1}$$
$$= c(a^n + a^{n+1}) + d(b^n + b^{n+1}) = c \cdot a^{n+2} + d \cdot b^{n+2} = F_{n+2}$$
$$x^2 = x + 1 \rightarrow x^2 - x - 1 = 0 \rightarrow a = \frac{1+\sqrt{5}}{2} = \phi \, , b = \frac{1-\sqrt{5}}{2} = 1 - \phi$$

- Now we solve for the coefficients *c* and *d*

$$F_1 = 1 = \phi c + (1-\phi)d \, , F_2 = 1 = \phi^2 c + (1-\phi)^2 d$$
$$c = \frac{1}{2\phi - 1} = \frac{1}{\sqrt{5}} \, , d = \frac{1}{1 - 2\phi} = -\frac{1}{\sqrt{5}} \rightarrow F_n = \frac{1}{\sqrt{5}}[\phi^n - (1-\phi)^n]$$

- Since the second term vanishes as *n* increases without bound, the ratio of consecutive terms approaches *a*

Software Estimating Data Flow

- In a preferred detailed Software Cost Estimating / Inputs Risk scenario, each component is modeled separately, with data-driven uncertainty