Monte Carlo and probability bounds in R with hardly any data

Scott Ferson, Institute for Risk and Uncertainty, University of Liverpool

Society for Risk Analysis SRA Annual Meeting

Shortcourse held at the Society for Risk Analysis annual meeting at Crystal Gateway Marriott, Arlington, Virginia

10 or 14 December 2017

Presenter:
Scott Ferson
Institute for Risk and Uncertainty
University of Liverpool
https://sites.google.com/site/scottfersonsite/

Outline

- Getting started in R
- Monte Carlo simulation
- Probability bounds analysis
- Psychometry of uncertainty
- Correlations and dependencies
- Model uncertainty
- Sensitivity analysis
- Backcalculation
- Interval statistics
- Computing with confidence
- Fermi estimates
- Imprecise probabilities
- Case studies (NASA and EPA)
- Conclusions

“Open question”

Suppose

A is in [2, 4]

B is in [3, 5]

What can be said about the sum $A+B$?

Getting started in R

Installing libraries we'll need

- Click Packages / Install package(s)…, select any CRAN mirror, and click OK
- After a pause, a list of packages will pop up
- Select the package MASS, and click OK
- You may wish to install other packages

Install the sra.r Monte Carlo library

- If you left R, invoke it again
- Enter **rm(list = ls())** to clear its memory
- Click File / Source R code on the main menu
- Locate and Open the file **sra.r**
- You'll get the message "`:sra> library loaded`"

SKIP

Testing the sra.r installation

- Start a new instance of R, and enter the command **rm(list = ls())**
- Click File / Source R code on the main menu; find and Open the file **sra.r**
- The R Console will display something like

```
> source("C:\\Users\\workshop sra\\sra.r")
:sra> library loaded
```

- Enter **plot(runif(100))** which will open a plot window with random points
- Click History / Recording on the main (if History is missing, click the plot)
- Click on the R Console, click File / Open script, and find and Open **test sra.r**
- Press Ctrl-A to select all its text, then press Ctrl-C to copy it to the clipboard, and close the R Editor window
- Click on the R Console, and press Ctrl-V to paste the text into the R Console for execution
- Click on the plot window, and press PgUp and PgDn to scroll through graphical results of the test calculations
- There should be no error messages on the R Console

sra.r Monte Carlo library

a = normal(5,1)

a

b = uniform(2,3)

b

a + b

Generalized convolutions

a * b

pmin(a,b)

c = a + b * beta(2,3) / a

mean(c)

SKIP

Functionality in sra.r

- Distributions

bernoulli, beta, betapert, binomial, cauchy, chi, chisquared, delta, dirac, discreteuniform, exponential, exponentialpower, extremevalue, F, fishersnedecor, fishertippett, fisk, frechet, gamma, gaussian, geometric, gev, generalizedpareto, gumbel, histogram, inversegamma, laplace, logistic, loglogistic, lognormal, logtriangular, loguniform, MEdiscretemean, MEdiscreteminmax, MEmeansd, MEmeanvar, MEminmax, MEminmaxmean, MEminmaxmeansd, MEminmaxmeanvar, MEminmean, MEquantiles, mc (pointlist), negativebinomial, normal, N, pareto, pascal, powerfunction, poisson, quantiles, rayleigh, reciprocal, shiftedloglogistic, skewnormal, student, trapezoidal, triangular, uniform, U, weibull

- Operators, functions and commands

+, *, -, /, ^, abs, acos, and, asin, atan, atan2, ceiling, complement, cor, cos, cut, edf, exp, fivenum, floor, iqr, IQR, left, lines, log, mean, median, mixture, negate, not, oppositeconv.mc, or, percentile, perfectconv.mc, plot, pmax, pmin, prob, quantile, random, reciprocate, right, round, samedistribution, sd, show, sign, sin, specific, sqrt, summary, tan, trunc, var

Monte Carlo simulation

Traditional practice

- Most assessments are deterministic
- Some are deliberately conservative, but
  - degree of conservatism is opaque, unquantified, and can be inconsistent
  - difficult to characterize risks, except in extreme situations

What’s needed

An assessment should tell us

- How likely the various consequences are
- How reliable the conclusions are

Probabilistic models

- The same as deterministic models except that point values are replaced by probability distributions

- Well, *almost*
  - Ensembles
  - Distribution shapes and tails
  - Dependencies
  - Backcalculation

What is probability?

- The frequentist view is that the probability of an event is its *frequency* given a long sequence of independent trials

- The subjectivist view is that probability of an event is the *degree of belief* that a person has that the event will occur

- *Don't confuse or mix these ideas*

SKIP

Bayesian probability

- Asserting *A* means you'll pay \$1 if not *A*
- If the probability of *A* is *P*, then a Bayesian agrees to assert *A* for a fee of \$(1-*P*), ***and*** assert not-*A* for a fee of \$*P*
- For Bayesians, belief implies action: bets
- People will have different *P*s for an *A*

SKIP

Frequentist probabilty

- Let N_A be the number of times event A occurs out of N trials
- The probability of A is N_A/N, as $N \rightarrow \infty$
- If the probability of event A is P(A) then event A will occur N×P(A) times, on average, out of every N times

SKIP

Distribution

- Arrangement of a variable's values showing their commonness of occurrence
- Several ways to express commonness
  - Counts (most concrete conception)
  - Relative frequencies (finite populations)
  - Probabilities (more abstract)
- Probability distributions used to model *both* variability and incertitude

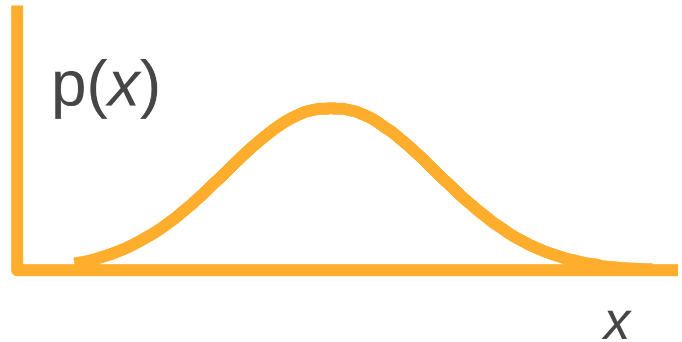
Probability distributions

- *Distribution*: bunch of values (ignoring order)
- *Stationary* distribution: unchanging shape
- *Random values* from a distribution are values chosen in an unpredictable order (unrelated to past and future choices) but according to the specified frequencies

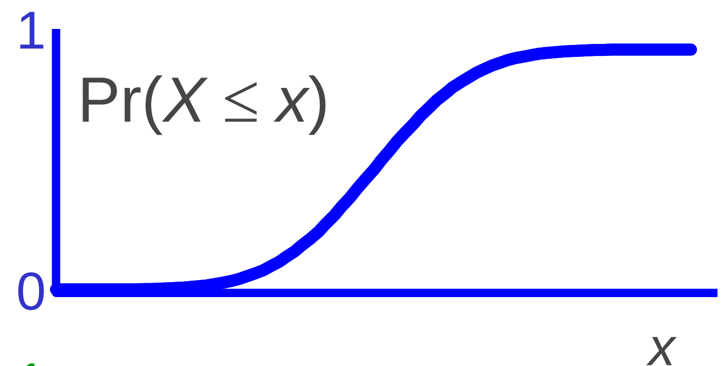
"random" means independent and from the same distribution

Probability distribution display

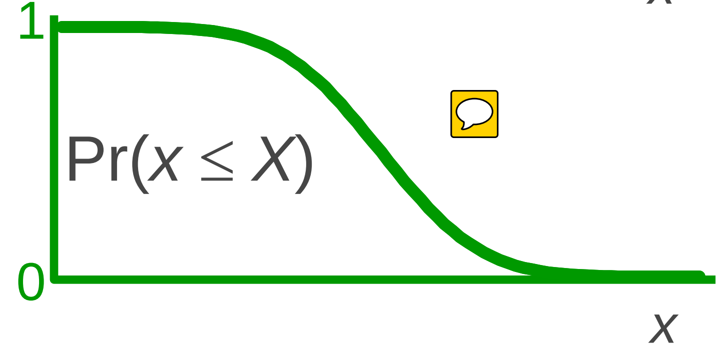
- Density distribution or mass distributions



- Cumulative distribution function (CDF, integral of density)

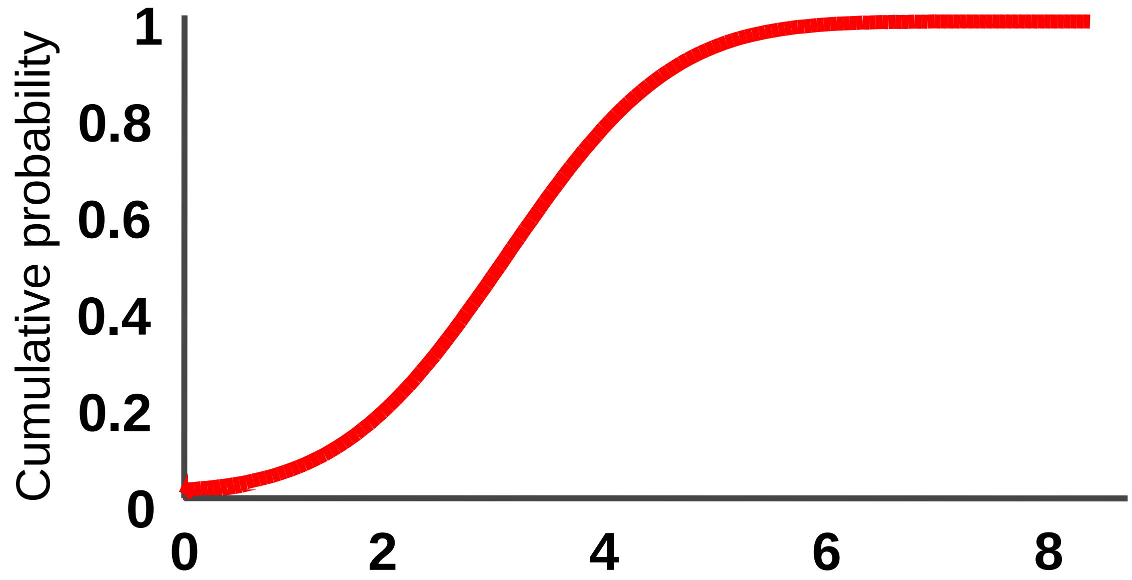


- Complementary CDF (exceedance risk)



Continuous distributions

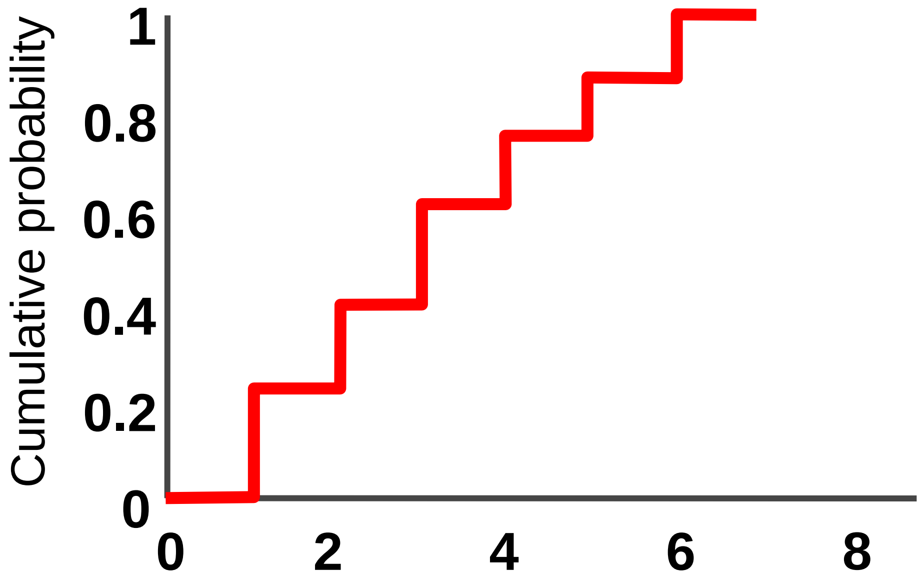
- normal
- uniform
- exponential
- lognormal
- triangular
- Weibull
- beta
- Laplace
- many, many more



Values less than 0, or larger than 6, are impossible. Most values are around 3. 80% of values are less than 4.

Discrete distributions

- binominal
- Poisson
- discrete uniform
- reciprocal



Distributions can also be “mixed” (neither continuous nor discrete)

Notation and language

- A "deviate" or "realization" is a *particular* real value of a random variable or a value drawn from a distribution
- U(a, b) denotes a uniform distribution with minimum a and maximum b, or sometimes a deviate from it
- N(μ, σ) denotes a normal distribution with mean μ and standard deviation σ, or a deviate from it
- A tilda ~ is read "is distributed as"
- A distribution that describes the variation of a random variable is sometimes called its "law", so if $X \sim$ N(0,1) the law for X is normal

Ensemble

- Statisticians call it the “population” or “reference class”
- E.g., dysfunction rates in prostate patients
    - 50% of attempts at sex fail, or
    - 50% of men are totally impotent

Specification of ensembles

- Above all, be clear
  - 1% of lethal dose per day / lethal dose on 1% of days
  - Assessment units: birds, meals, days, fields, …
  - Assessment ensemble: all birds, or just those on site?
- Mismatched input distributions yield gibberish
- If the analyst can't tell you what ensemble the distribution represents, the analysis is probably meaningless

Computational methods

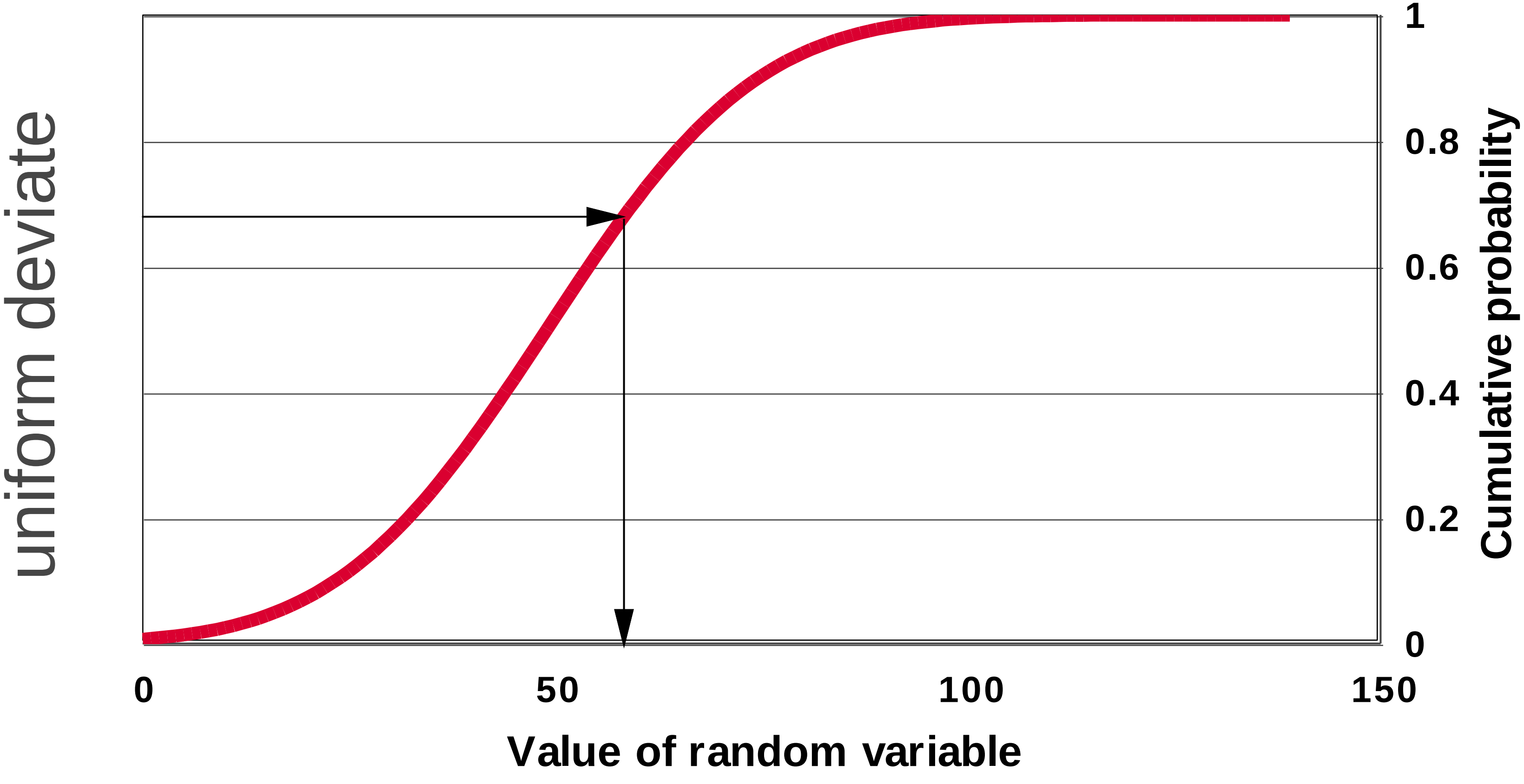
- Analytical approaches
  - Laplace and Mellin transforms
  - Delta method (just means and variances)
- Discrete probability distributions
- Monte Carlo simulation
  - Sampling values from input distributions
  - Computing function of inputs
  - Accumulating output histograms

Pseudorandom numbers

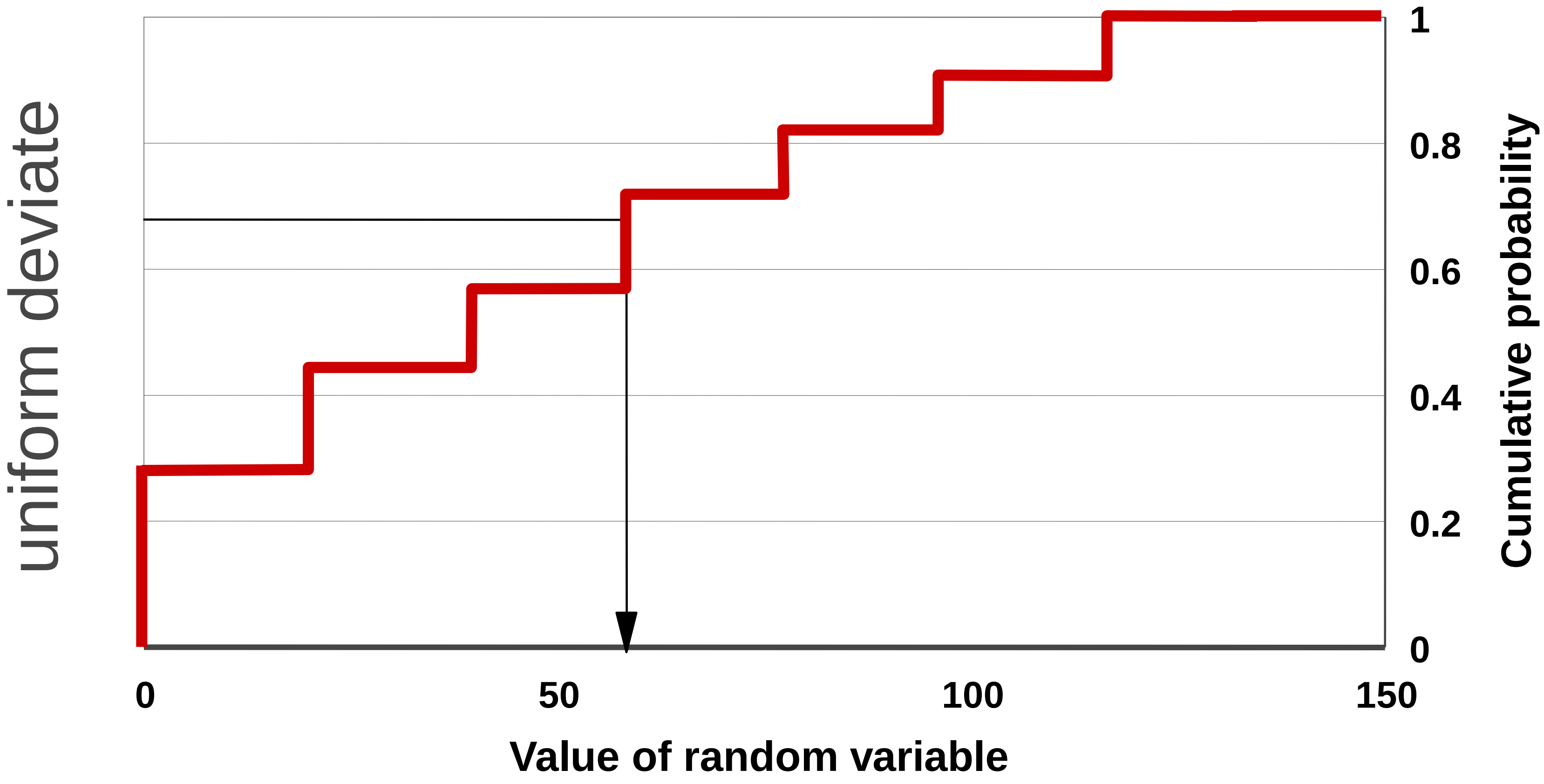
Sampling from a uniform

- Computerized eenie-meanie-miney-moe
  - Neumann's middle-square method
  - Linear congruential generators
  - Mersenne twister (bulkier but faster and better)
  - Blum-Blum-Shub
- Smart to repeat simulations with different generators
- The quantity $a + (b - a)u$ varies uniformly over $[a, b]$ if u varies uniformly over $[0,1]$

Sampling continuous distributions



Sampling discrete distributions



SKIP

Generating random deviates

- *Inverse transform sampling* can make realizations from an *arbitrary* distribution using uniform deviates

- *Acceptance-rejection method* can be used when the CDF has no closed form

- Many distribution shapes have special algorithms that are faster or better

SKIP

Box-Muller algorithm

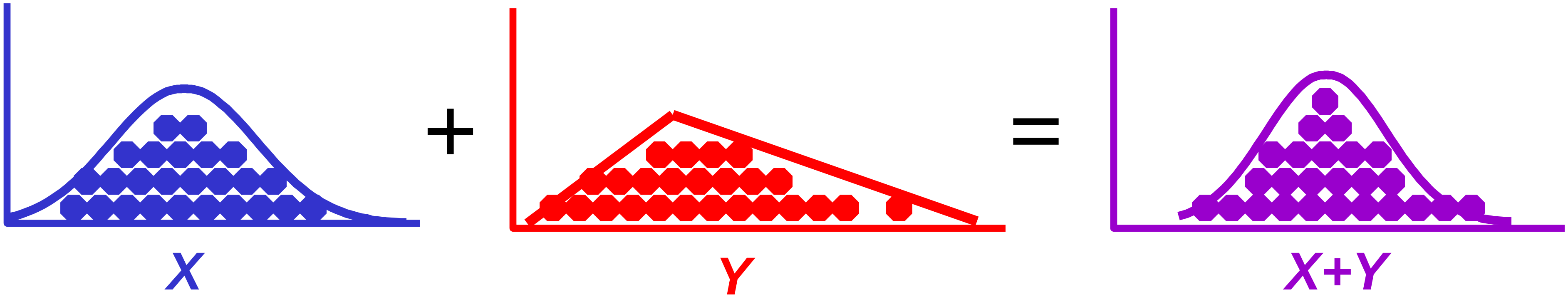
$$N(0,1) \sim \sqrt{-2 \ln(u) \times \sin(2\pi\, v)}$$

where u, $v \sim U(0,1)$, and $u \perp v$

- Specialized generator for normal deviates
- Faster than inverting the cumulative normal distribution which has no closed form
- General deviates $N(\mu, \sigma) \sim \sigma\, N(0,1) + \mu$

Convolution
(arithmetic on distributions)

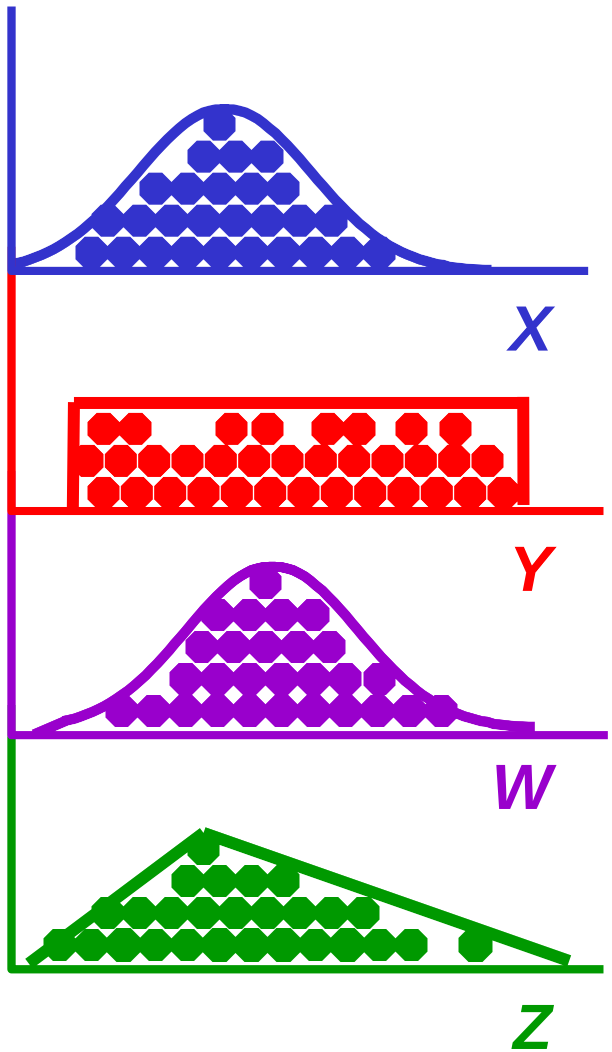
Convolution

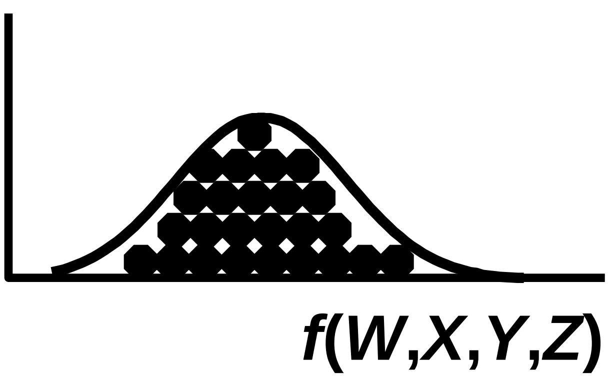


Convolution is getting the output distribution from the input distributions. Monte Carlo is one way to do this.

Integrated calculations

Sample realizations from each input distribution, then combine them all to compute a (scalar) output value. Repeat many times to accumulate a distribution of output values.





This is unlike interval analysis which solved operations sequentially

Example: contaminant plume

Lobascio

- Hydrocarbon in groundwater near some wells
- Constant, one-dimensional, uniform Darcian flow
- Homogeneous properties (e.g., no pipes, conduits, barriers or differential permeability among layers)
- Linear retardation
- No dispersion

- How long before the contaminant reaches the wells?

Example: travel time

$$T = \frac{(n + BD \times foc \times Koc)L}{K \times i}$$

Parameter	Units	Min	Max	Mean	Stdv	Shape
L source-receptor distance	m	80	120	100	11.55	uniform
i hydraulic gradient	m/m	0.0003	0.0008	0.00055	0.0001443	uniform
K hydraulic conductivity	m/yr	300	3000	1000	750	lognormal
n effective soil porosity	-	0.2	0.35	0.25	0.05	lognormal
BD soil bulk density	kg/m^3	1500	1750	1650	100	lognormal
foc fraction organic carbon	-	0.0001	0.005	0.00255	0.001415	uniform
Koc organic partition coefficient	m^3/kg	5	20	10	3	normal

Marco Lobascio in *EnviroNews*

Implementation

L = uniform(80, 120) # [m], source-receptor distance

i = uniform(0.0003, 0.0008) # [], hydraulic gradient

K = lognormal(1000, 750) # [m yr^{-1}], hydraulic conductivity

n = lognormal(0.25, 0.05) # [], effective soil porosity

BD= lognormal(1650, 100) # [kg per m^3], soil bulk density

foc = uniform(0.0001, 0.005) # fraction organic carbon

Koc = normal(10, 3) # [m^3 per kg], partition coefficient

T = (n + BD * foc * Koc) * L / (K * i) # all variables assumed independent

summary(T)

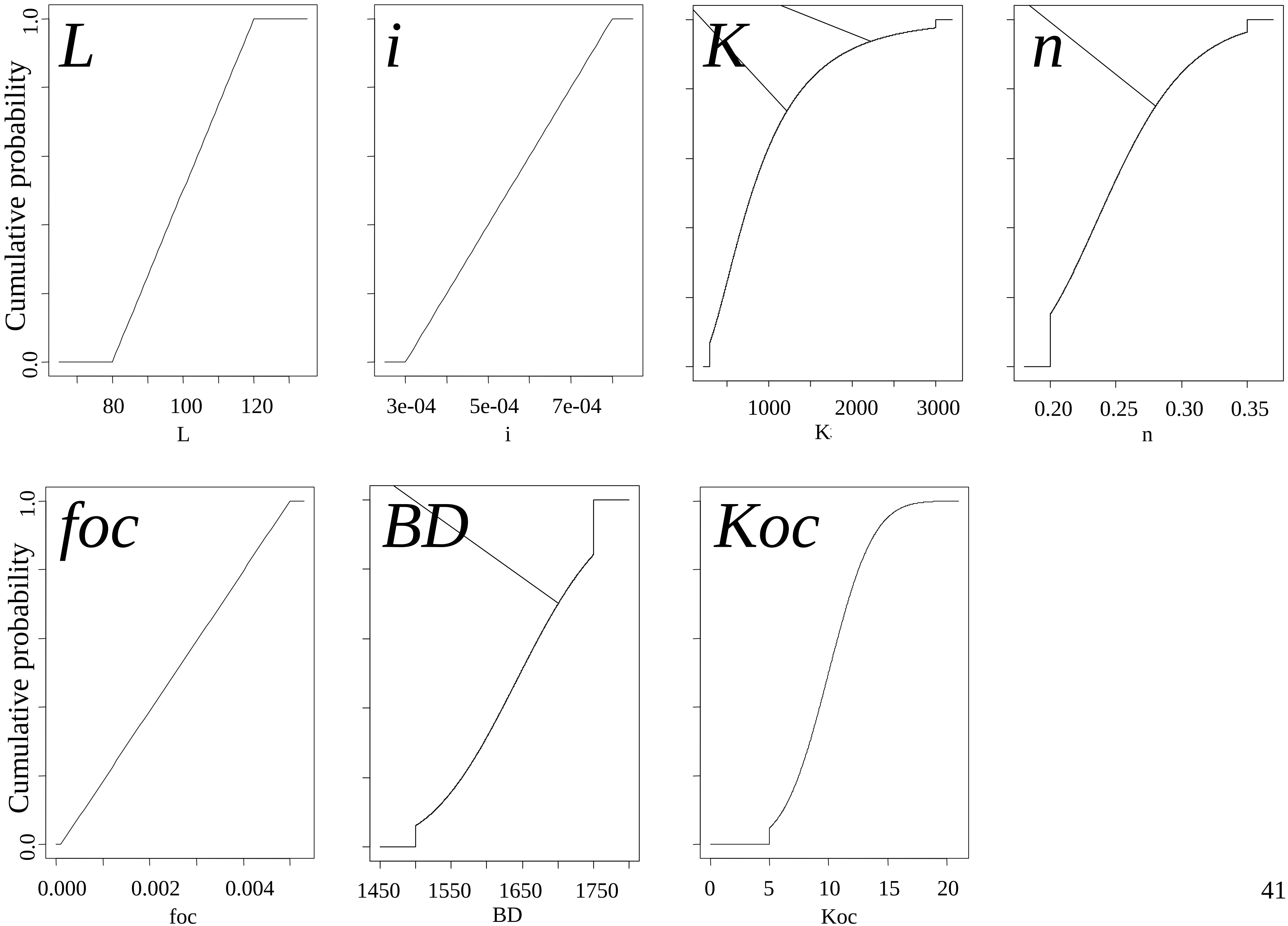
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-4038	3598	8040	12910	16190	471200

How'd we get negative some travel times?

Implementation

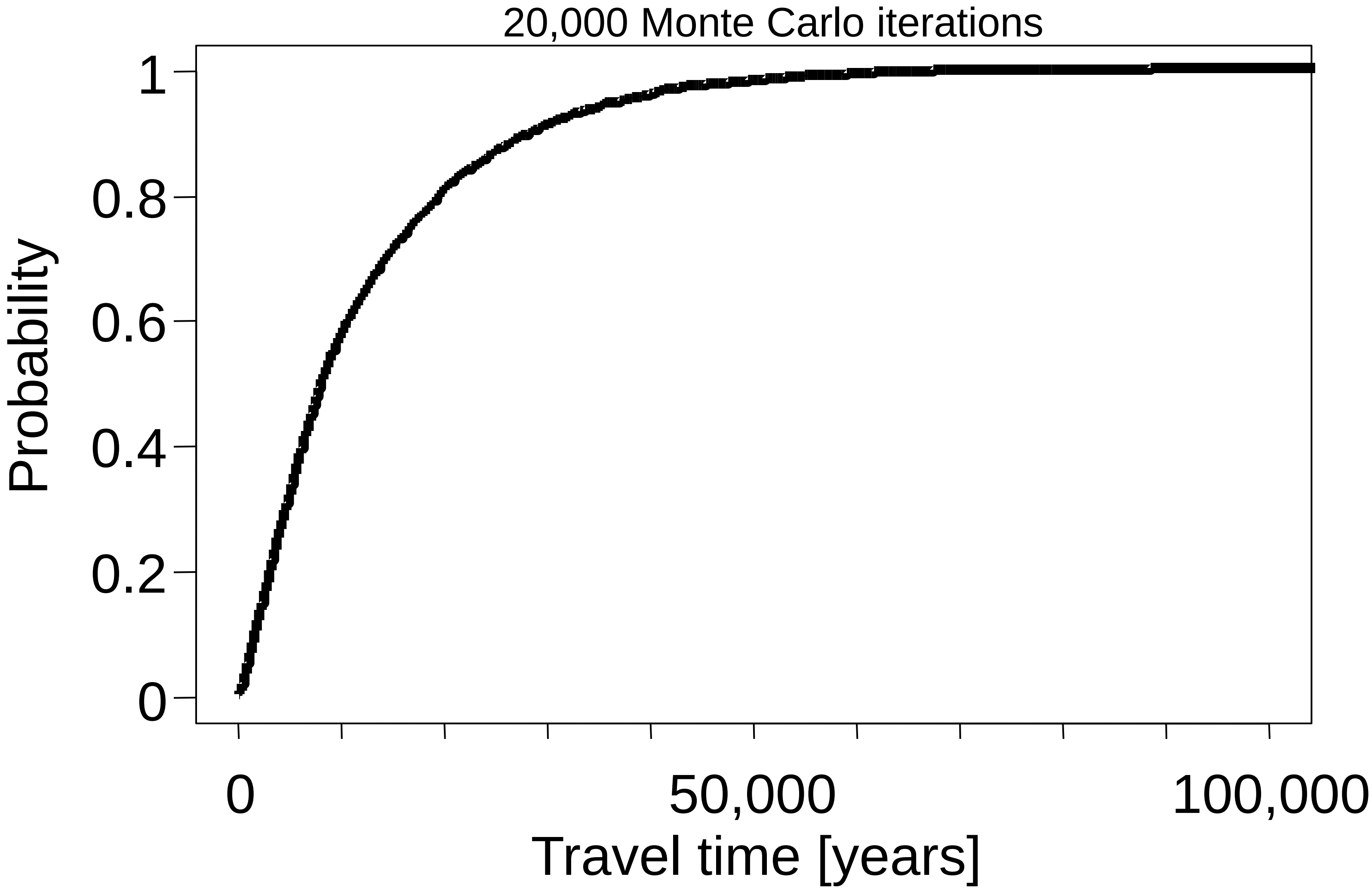
```
L   = uniform(80, 120)                # [m], source-receptor distance
i   = uniform(0.0003, 0.0008)         # [], hydraulic gradient
K   = lognormal(1000, 750)            # [m yr⁻¹], hydraulic conductivity
K   = truncate(K, 300, 3000)
n   = lognormal(0.25, 0.05)           # [], effective soil porosity
n   = truncate(n, 0.2, 0.35)
BD= lognormal(1650, 100)              # [kg per m³], soil bulk density
BD= truncate(BD, 1500, 1750)
foc = uniform(0.0001, 0.005)          # fraction organic carbon
Koc = normal(10, 3)                   # [m³ per kg], partition coefficient
Koc = truncate(Koc, 5, 20)

T  = (n + BD * foc * Koc) * L / (K * i)
summary(T)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
87	3693	8070	12200	16230	135900



L
i
K
n
foc
BD
Koc
Cumulative probability
Cumulative probability
0.0
1.0
0.0
1.0
80
100
120
L
3e-04
5e-04
7e-04
i
1000
2000
3000
K
0.20
0.25
0.30
0.35
n
0.000
0.002
0.004
foc
1450
1550
1650
1750
BD
0
5
10
15
20
Koc

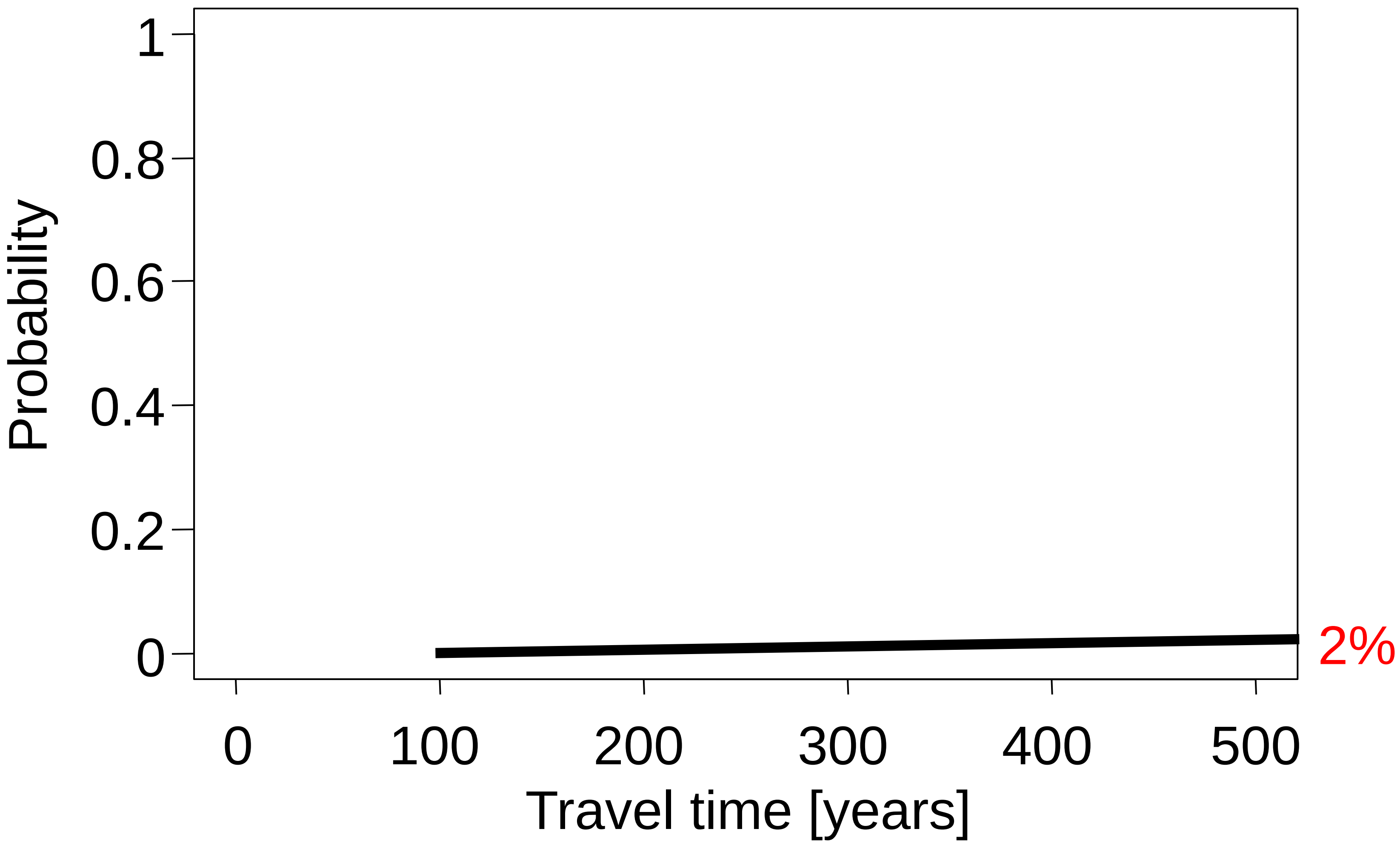
Result of the simulation



Interpreting the result

- What is the minimum time? Maximum?
    - » 87 years, 136,000 years
- What happens when reps are increased?
    - » Smaller min, larger max, approaching interval range
- Why is the maximum time so much smaller than the upper limit which is 234,000 years?
    - » Distribution highly skewed
- If the random variables weren't truncated, what would the min and max be?
    - » Zero or even *negative* travel times, larger max time

Detail


Probability
1
0.8
0.6
0.4
0.2
0
2%
0
100
200
300
400
500
Travel time [years]

Selecting input distributions

How to specify distributions

- Default distribution
- PERT
- Expert opinion
- Maximum entropy
- Fitted distribution (goodness of fit)
  - Method of matching moments
  - Maximum likelihood
- Empirical distribution

No data

Lots of data

Default distributions

- Physicists overuse normal distributions
- Risk analysts overuse uniform distributions
- Reliability engineers overuse Weibulls
- Ecologists overuse lognormals
- Statisticians overuse beta distributions
- Food safety overuses triangular distributions

Cooper's quip

Arguments (frequently abused)

- Sum of many independent comparable terms (normal)
- Product of many independent factors (lognormal)
- Largest of many values (Gumbel, Weibull, Fréchet)
- Lifetimes with neither aging nor burn-in (exponential)
- Symmetry (uniform)
- Rare discrete events (Poisson)
- Independent trials (binomial)

Fitting distributions

- Makes use of available data
- Assumes you know the distribution family
- Different methods give different results
  - Method of moments
  - Maximum likelihood
  - Regression methods

Method of matching moments

- Pick the distribution family
  - Mechanistic argument or mathematical convenience
- Compute moments of the data
- Select distribution from family with same moments
  - E.g., exponential (1), normal (2), triangular (3)

- Other issues
  - Check the goodness of fit of the answer
  - Software can check many shapes at once
  - MoM not always available, lacks optimality properties

Example

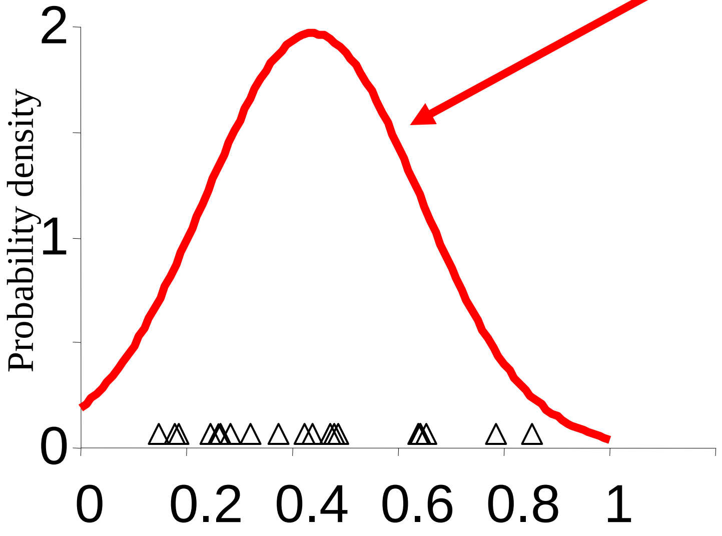
Data

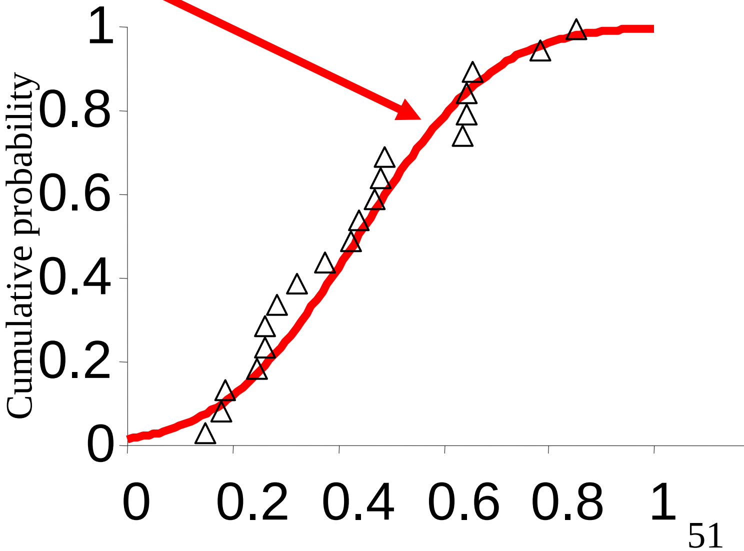
0.653
0.178
0.263
0.424
0.284
0.438
0.471
0.852
0.480
0.375
0.148
0.185
0.320
0.642
0.247
0.784
0.643
0.261
0.636
0.487

mean(Data) = 0.439

sd(Data) = 0.202655

normal(0.439, 0.203)

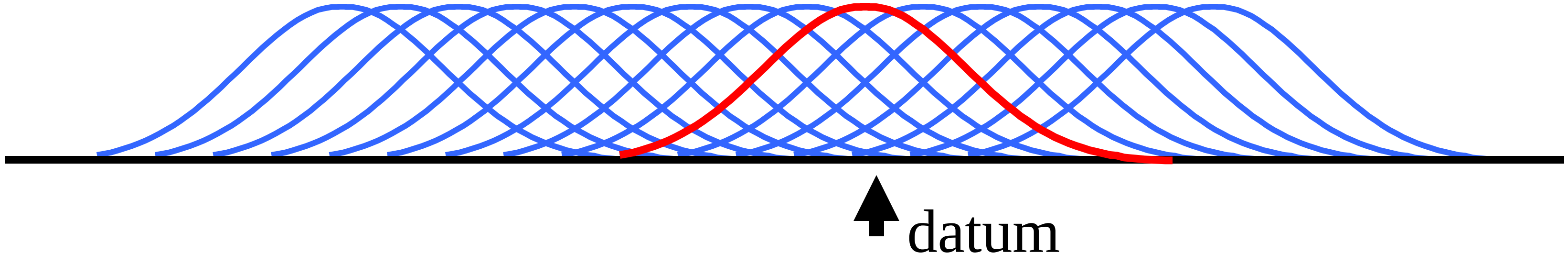




SKIP

Maximum likelihood (ML)

- Find the distribution that maximizes the likelihood of having observed the data
- Likelihood is the probability of seeing the data assuming the distribution actually has some form



SKIP

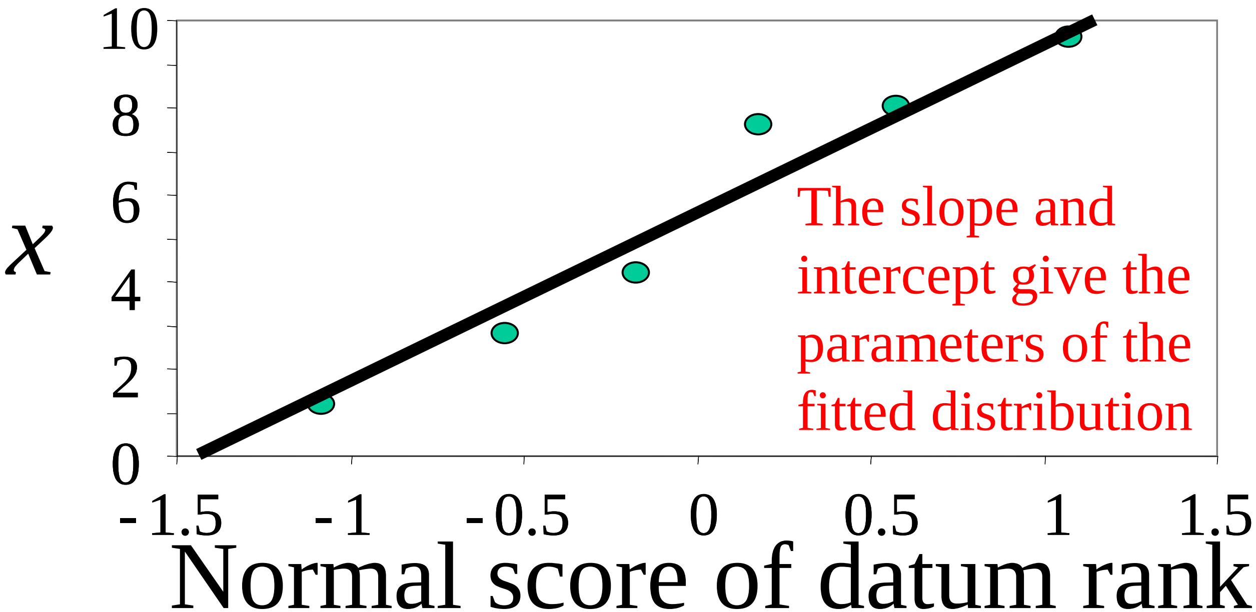
Maximum likelihood in practice

- When data are sampled independently, just multiply the likelihoods
- Formulas for many shapes known
- But they can sometimes be scary
- Maximum likelihood agrees with MoM for some distributions (e.g., normal and exponential)

SKIP

Graphical/regression methods

- Rank the data
- Transform ranks by probability distribution
- Regress data against the transformed ranks



SKIP

Goodness of fit (GoF)

- Goodness of fit measures
  - Chi squared
  - Anderson-Darling

focuses on tails

  - Kolmogorov-Smirnov
  - Specialized tests for particular families

    (N,L : Shapiro-Wilks, Shapiro-Frances, D'Agostino, Lillefors)

- EPA says it's less important than sensibleness
  - Recommend bypassing a better fitting distribution for one that is plausible under some *mechanism*

Empirical distributions

- Distribution characterized by *observed* data

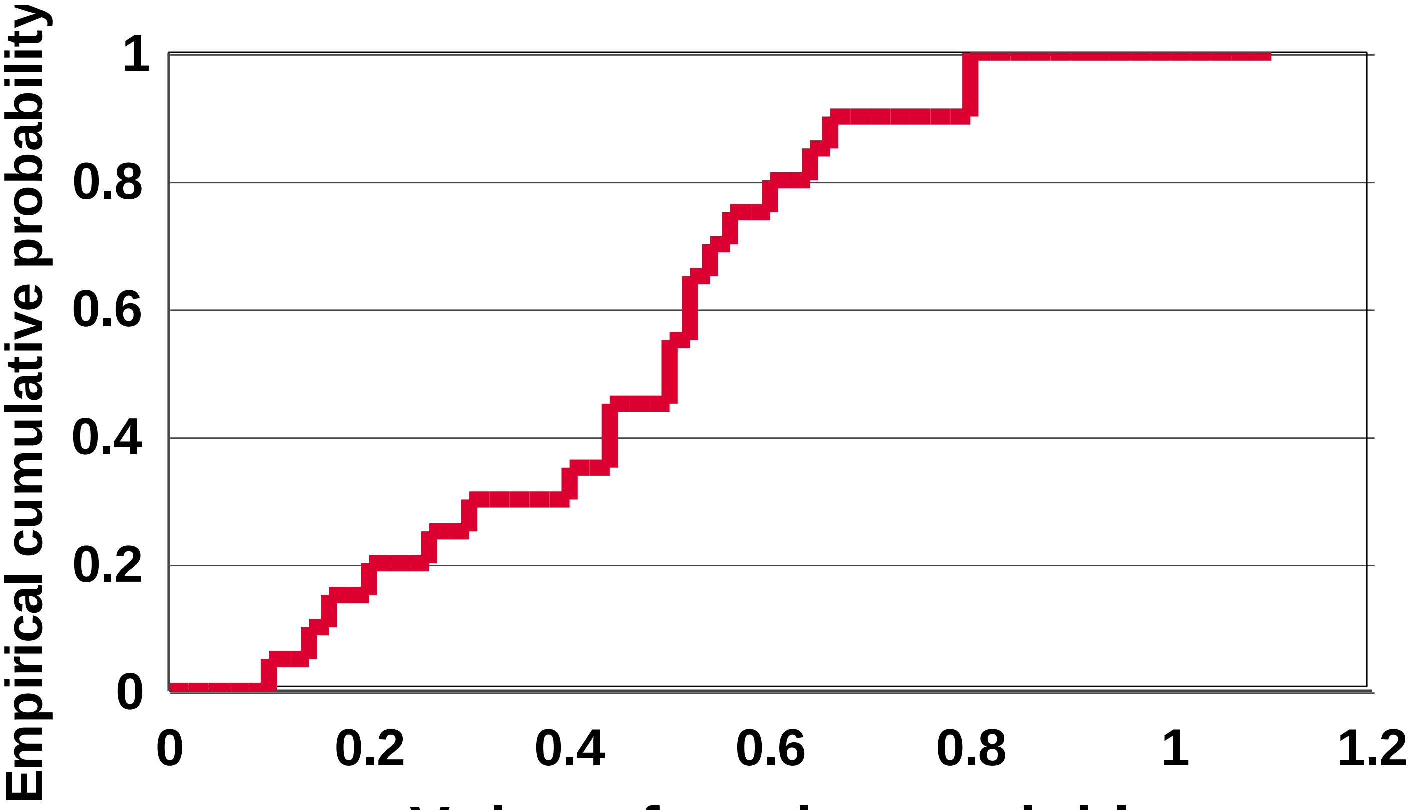
$$F(x) = {}^{1}/_{N} \Sigma \, \#(X_i < x)$$

- Assumes data were randomly collected
- Very useful when data are abundant
- *Doesn't need a shape assumption*

Example

Data

0.653
0.178
0.263
0.424
0.284
0.438
0.471
0.852
0.480
0.375
0.148
0.185
0.320
0.642
0.247
0.784
0.643
0.261
0.636
0.487



Drawbacks of EDFs

- Need a fair amount of data to yield reliable estimates of distributions
- Tails are likely to be poorly modeled
- Data usually underestimate the range
- Depends on "plotting position"

Selecting distributions *without* data

- PERT
- Maximum entropy
- Fermi estimates
- Expert elicitation

PERT

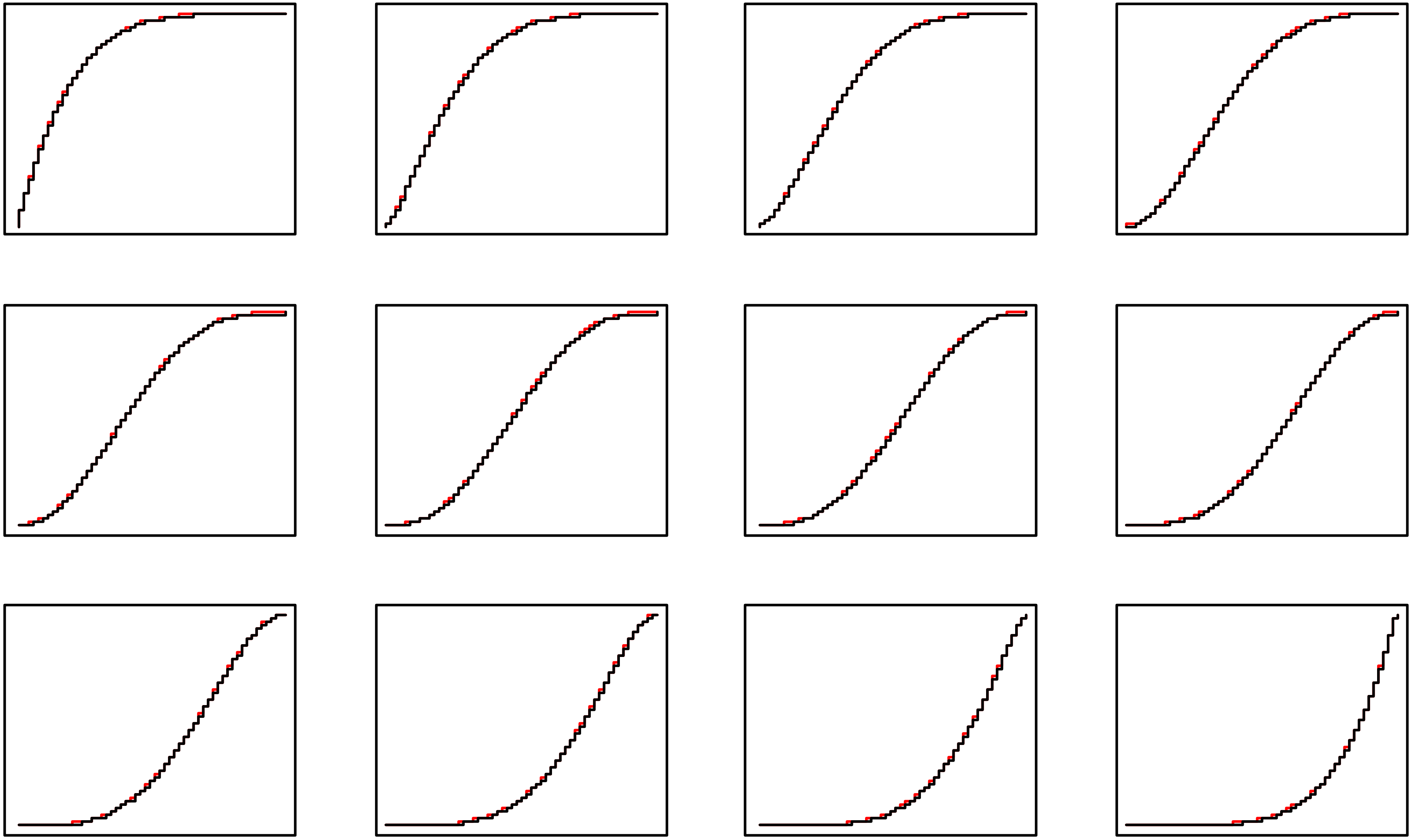
Program Evaluation and Review Technique

- Originally for costing and scheduling but now common
  - Developed in a "few weeks" by U.S. Navy contractors
- Uses expert opinion {min, max, mode} for each input

BetaPERT is $a + (b - a)\text{beta}(\alpha_1, \alpha_2)$, where a = min, b = max, c = most likely value, $\alpha_1 = (\mu - a)(2c - a - b)/[(c - \mu)(b - a)]$, $\alpha_2 = \alpha_1(b - \mu)/(\mu - a)$, $\mu = (a + b + 4c)/6$

- Can specify other distributions if desired
  - Primavera's PertMaster knows 10 shapes, including "custom"
- Don't need to know correlations or dependence
- Usually implemented in a Monte Carlo simulation

Every betaPERT looks like one of these



Example: explosion from gas leaks

Arthur Valais posted this example on the Riskanal listserver

- Min, max and mode for two gas line leaks
  - Propane: {1, 10, 2.5} L per minute
  - Butane: {4, 12, 9.75} L per minute
- Room volume $10 \times 8 \times 5$ m^3
- Duration 72 hours
- What's the chance the gas is in its explosive concentration range (between 2.2% and 8.4%)
- Could PERT's answer be considered reliable?

Advantages of PERT

- Widely used in some quarters
- Can use expert opinion when data are sparse
- Specifies distributions from three parameters
- Extra information about shapes can be used
- Doesn't need correlations or dependencies
- Outputs are full distributions

Disadvantages of PERT

- Can’t use information about correlations
- Can’t account for general dependencies
- Input shapes are totally arbitrary
- Outputs may not approximate reality
- Not mathematically well justified

Maximum entropy

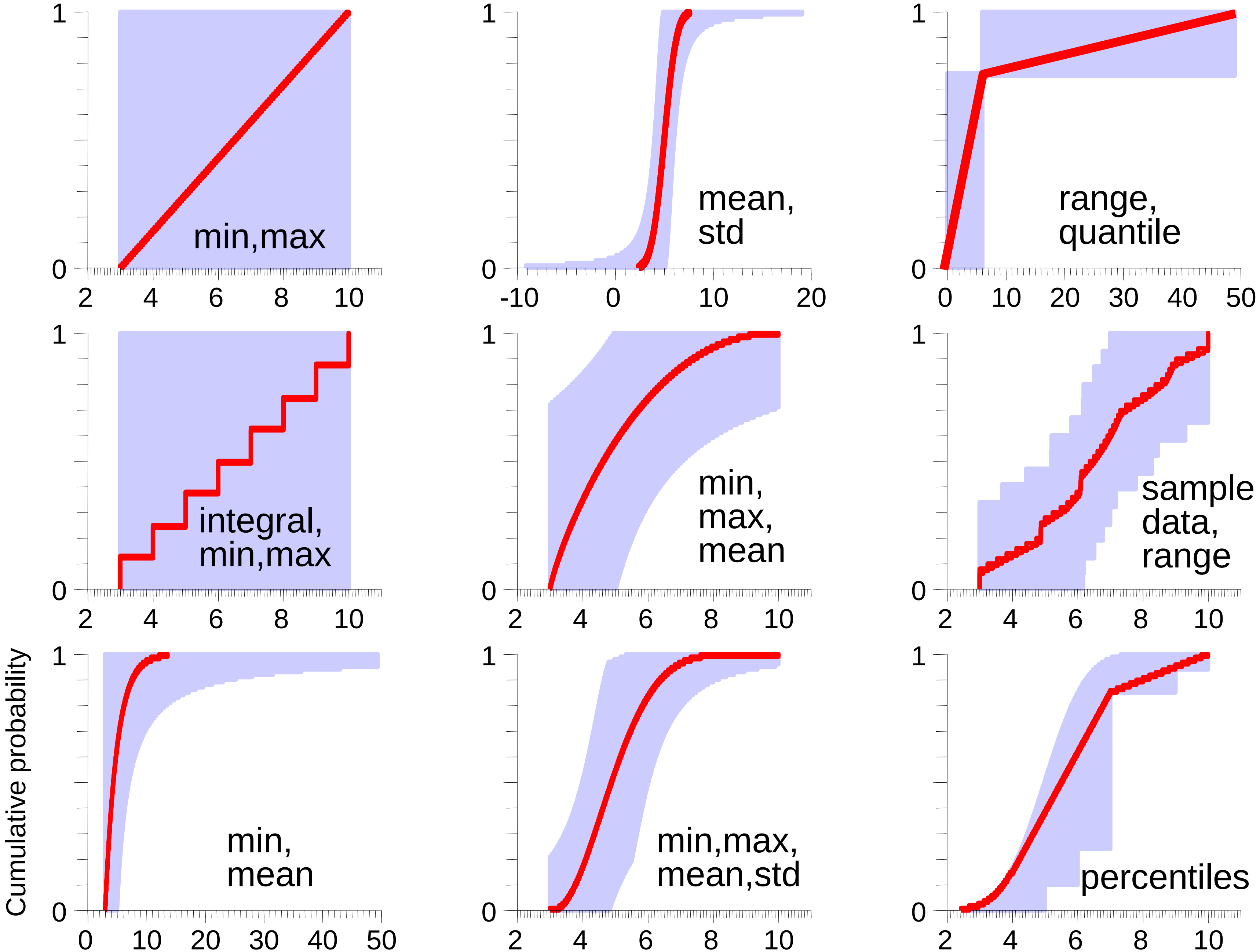
Maximum entropy

$$-\sum_{i=1}^{n} p_i \ln p_i$$

- Information theory
- All possibilities are equiprobable
- Specified by known constraints (not data)
- Laplace's "principle of insufficient reason"
- Mathematically more defensible than arbitrary assignment of distributions
- Much cheaper than expert elicitation

Maximum entropy solutions

When you know	*Use this shape*
{minimum, maximum}	uniform
{mean, standard deviation}	normal
{minimum, mean}	exponential
{min, max, some quantiles}	piecewise uniform
{minimum, maximum, mean}	Sawin-Conrad
{mean, geometric mean}	gamma
{minimum, maximum, mode}	?
{min, max, mean, stddev}	?



min,max
mean,
std
range,
quantile
integral,
min,max
min,
max,
mean
sample
data,
range
min,
mean
min,max,
mean,std
percentiles
Cumulative probability

Example: groundwater travel time

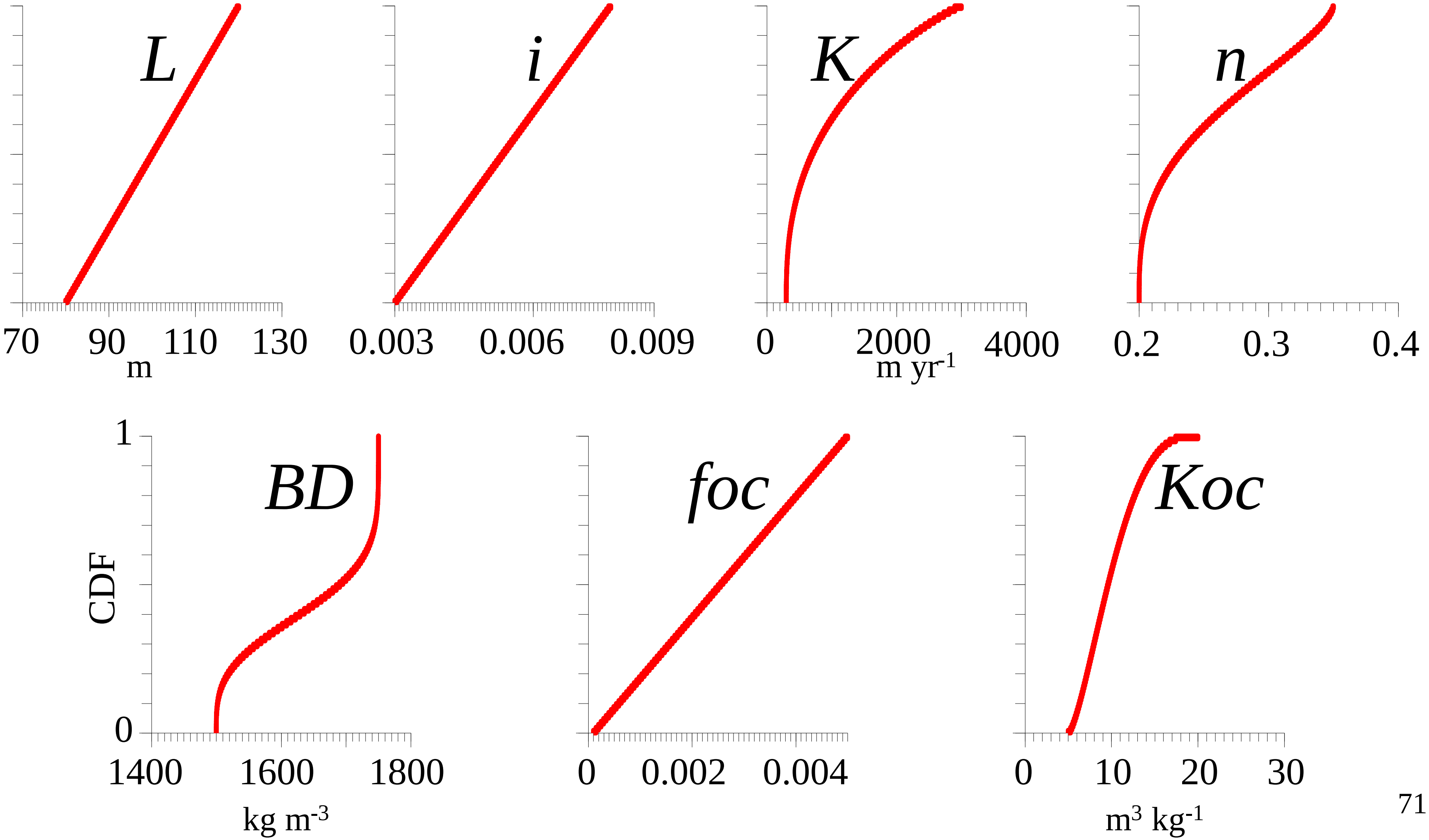
$$T = \frac{\left(n + BD \times foc \times Koc\right)L}{K \times i}$$

Parameter		Units	Min	Max	Mean	Stdv
L	source-receptor distance	m	80	120	100	11.55
i	hydraulic gradient	m/m	0.0003	0.0008	0.00055	0.0001443
K	hydraulic conductivity	m/yr	300	3000	1000	750
n	effective soil porosity	-	0.2	0.35	0.25	0.05
BD	soil bulk density	kg/m^3	1500	1750	1650	100
foc	fraction organic carbon	-	0.0001	0.005	0.00255	0.001415
Koc	organic partition coefficient	m^3/kg	5	20	10	3

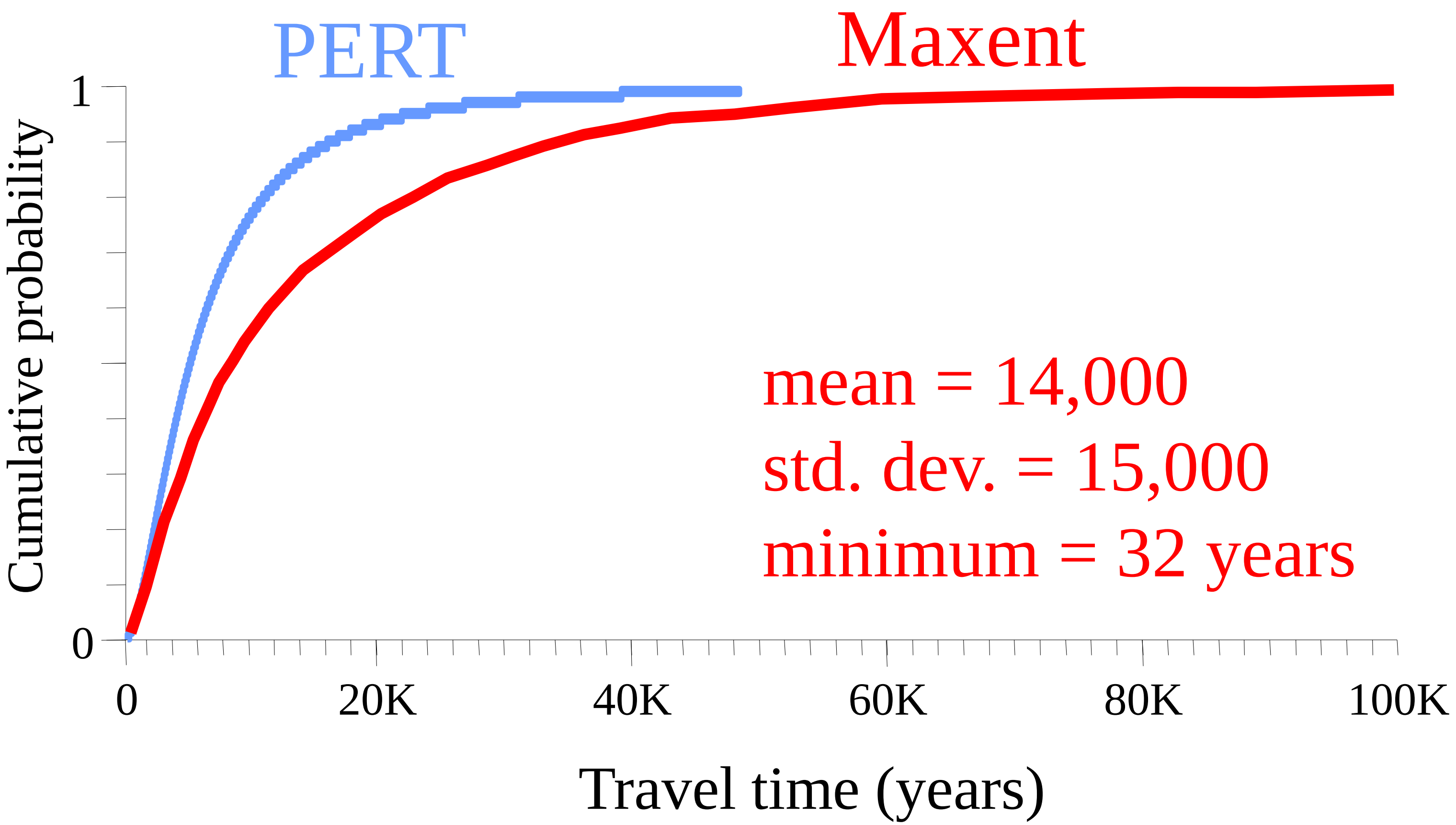
Implementation

```
L   = MEmmms(80,120,100,11.55)                      # source-receptor distance
i   = MEmmms(0.0003,0.0008,0.00055,0.0001443)       # hydraulic gradient
K   = MEmmms(300,3000,1000,750)                     # hydraulic conductivity
n   = MEmmms(0.2,0.35,0.25,0.05)                    # effective soil porosity
BD  = MEmmms(1500,1750,1650,100)                    # soil bulk density
foc = MEmmms(0.0001,0.005,0.00255,0.001415)         # fraction organic carbon
Koc = MEmmms(5,20,10,3)                             # organic partition coefficient

Tind  = (n + BD * foc* Koc) * L / (K * i)
Tind
```

Maximum entropy inputs



Travel time


PERT
Maxent
1
0
Cumulative probability
0
20K
40K
60K
80K
100K
Travel time (years)
mean = 14,000
std. dev. = 15,000
minimum = 32 years

Example: mercury in wild mink

Location: Bayou d'Inde, Louisiana

Receptor: generic piscivorous small mammal

Contaminant: mercury

Exposure route: diet (fish and invertebrates)

Based loosely on the assessment described in "Appendix I2: Assessment of Risks to Piscivorus [sic] Mammals in the Calcasieu Estuary", *Calcasieu Estuary Remedial Investigation/Feasibility Study (RI/FS): Baseline Ecological Risk Assessment (BERA),* prepared October 2002 for the U.S. Environmental Protection Agency. See http://www.epa.gov/earth1r6/6sf/pdffiles/appendixi2.pdf.

Total daily intake from diet

$$TDI = FMR \times \left(\frac{C_{\text{fish}} \times P_{\text{fish}}}{AE_{\text{fish}} \times GE_{\text{fish}}} + \frac{C_{\text{inverts}} \times P_{\text{inverts}}}{AE_{\text{inverts}} \times GE_{\text{inverts}}} \right)$$

C_{fish}	mercury concentration in fish tissue
C_{inverts}	mercury concentration in invertebrate tissue
P_{fish}	proportion of fish in the mink's diet
P_{inverts}	proportion of invertebrates in the mink's diet
GE_{fish}	gross energy of fish tissue
GE_{inverts}	gross energy of invertebrate tissue
AE_{fish}	assimilation efficiency of dietary fish in the mink
AE_{inverts}	assimilation efficiency of dietary invertebrates in the mink
FMR	normalized free metabolic rate of the mink

What's known about the inputs

BW, Empirically well studied
Normal with known mean and dispersion

normal

GE_{fish}, GE_{inverts}, Many measurements
Normal with known mean and dispersion

normals

AE_{fish}, AE_{inverts}, Some field data
Mean and upper and lower values

maxent

C_{fish}, C_{inverts}, Dictated by EPA policy
95% upper confidence limit on concentration

scalars

P_{fish}, P_{inverts}, Assumed by the analyst
Constants

scalars

FMR, Regression on mink body mass
$FMR = a\ BW^{b}$, where *a* and *b* are intervals

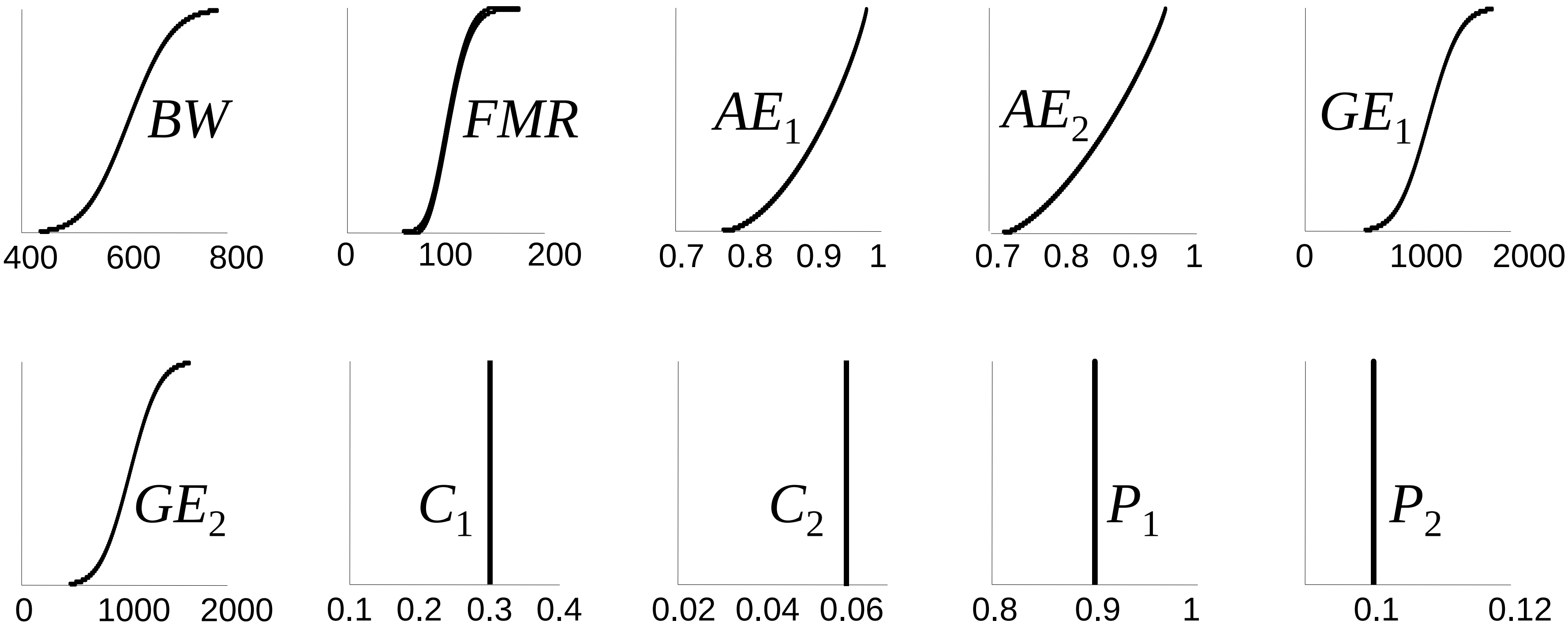
uniforms

Input assignments

BW = normal(608, 66.9) # gram
a = uniform(0.354,0.47)
b = uniform(0.836,0.888)
FMR = $a * BW$ ^ b # Kcal per kg per day
AE_{fish} = MEminmaxmean(0.77, 0.98, 0.91)
$AE_{inverts}$ = MEminmaxmean(0.72, 0.96, 0.87)
GE_{fish} = normal(1200, 240) # Kcal per kg
$GE_{inverts}$ = normal(1050, 225) # Kcal per kg
C_{fish} = 0.3 # mg per kg
$C_{inverts}$ = 0.06 # mg per kg
P_{fish} = 0.9
$P_{inverts}$ = 0.1

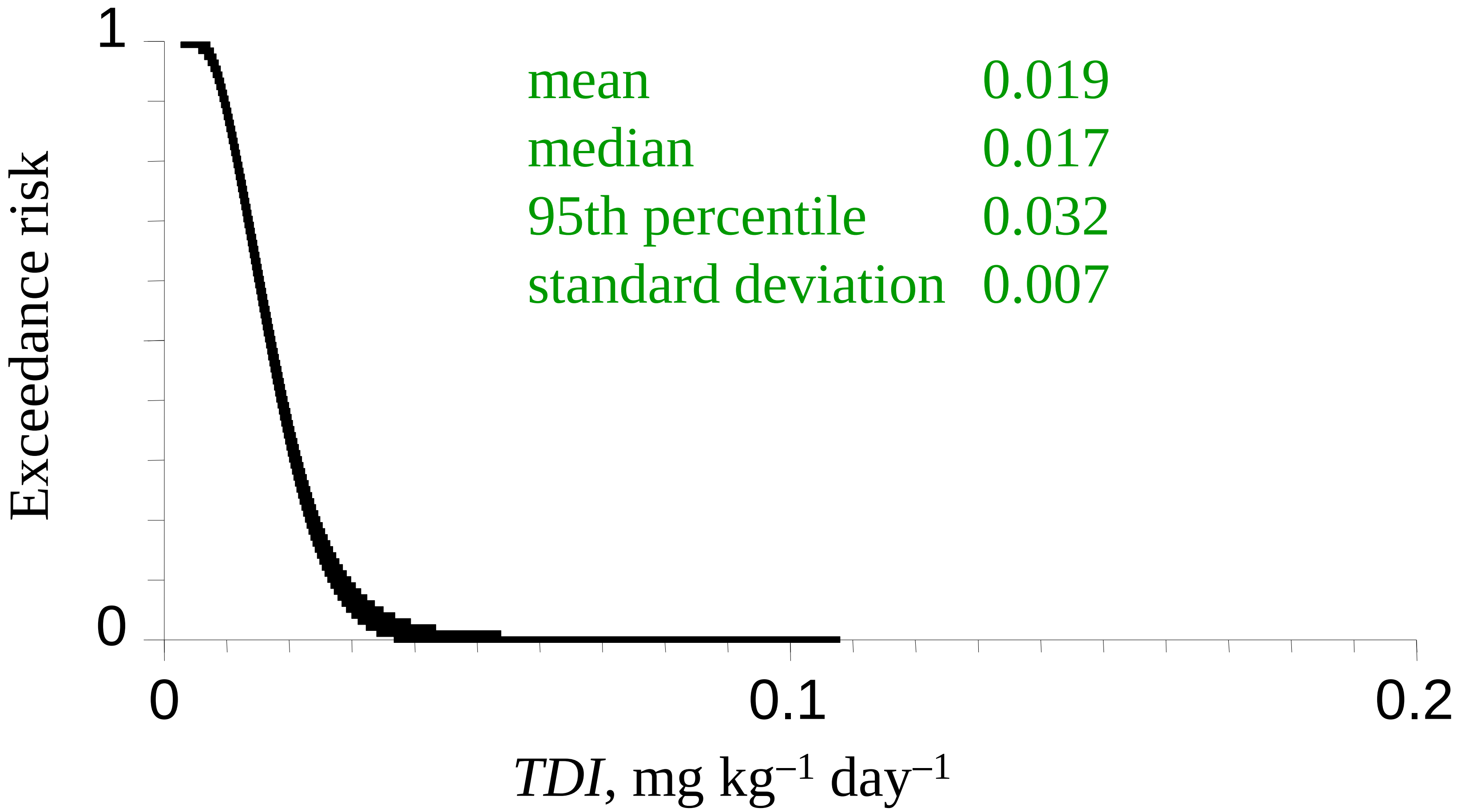
$$TDI=FMR*(C_{fish}*P_{fish}/(AE_{fish}*GE_{fish})+C_{invert}*P_{invert}/(AE_{invert}*GE_{invert}))$$

Input distributions



Subscript 1 denotes fish, 2 denotes inverts

Results


1
0
Exceedance risk
mean 0.019
median 0.017
95th percentile 0.032
standard deviation 0.007
0
0.1
0.2
TDI, mg kg⁻¹ day⁻¹

Maximum entropy's problem

- Depends on the choice of scale
- For instance, knowing the possible range for degradation rate yields one distribution, but knowing the possible range for half life yields an incompatible distribution even though the information is exactly the same

How to specify distributions

- Default distributions
  come right out of the book
- Fitted or empirical distributions
  usually not enough data available
- Extrapolations and surrogate data
  requires professional judgment
- Elicitation from experts
  expensive, controversial when experts disagree
- Maximum entropy criterion
  inconsistent through changes of scale

Many consider this the state of the art

Wrinkles

- Number of replications
- Multiple instantiations
- Dependence among variables
- Model uncertainty
- Backcalculation

How many replications

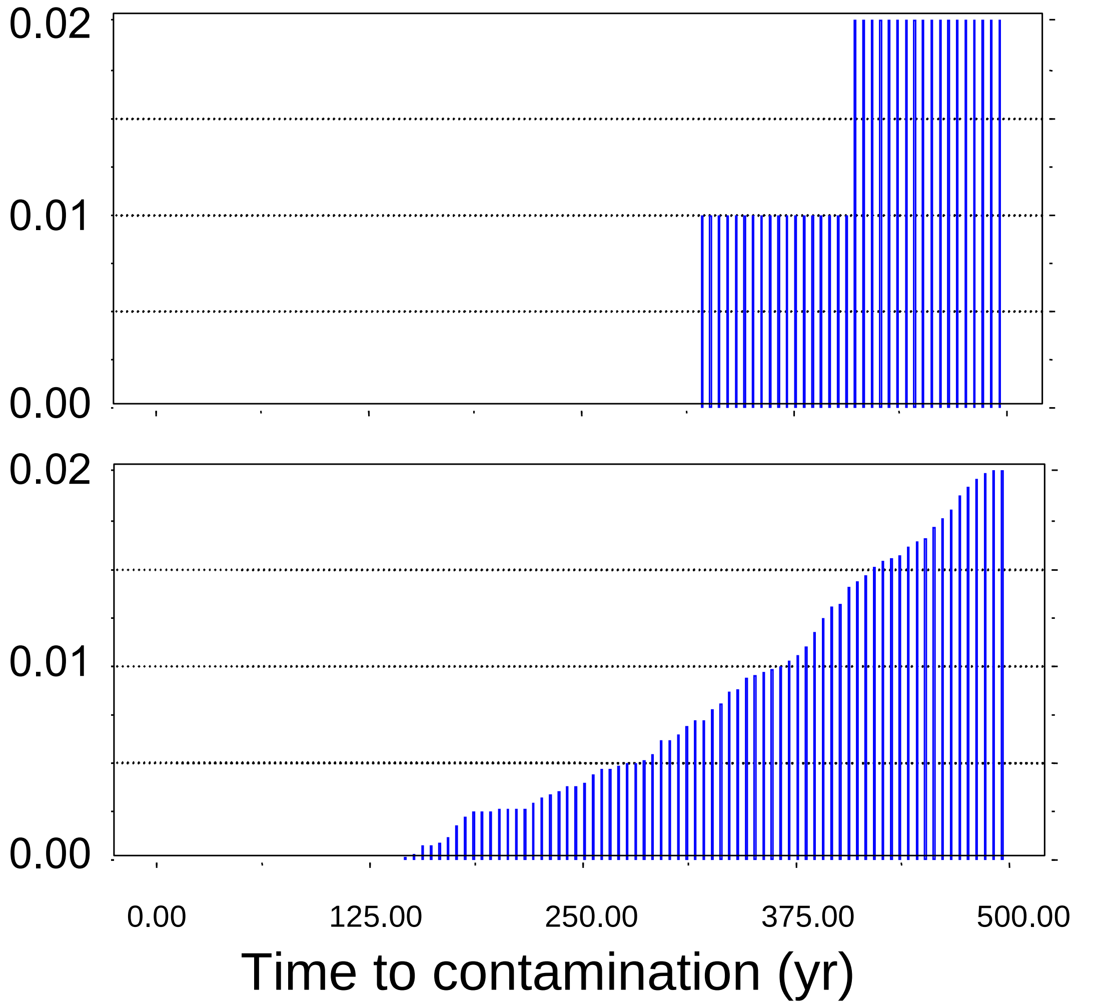
- More is always better
- Tails are especially hard to nail down
- Curse of dimensionality
- Latin hypercube sampling can help
- Repeat the simulation as a check ✓ **best**
- Confidence intervals on fractiles
- Kolmogorov-Smirnov limits on distributions

SKIP

100 versus 10,000 replicates

100 Trials: close-up of left hand tail

10,000 Trials: close-up of left hand tail



SKIP

Confidence interval for a fractile

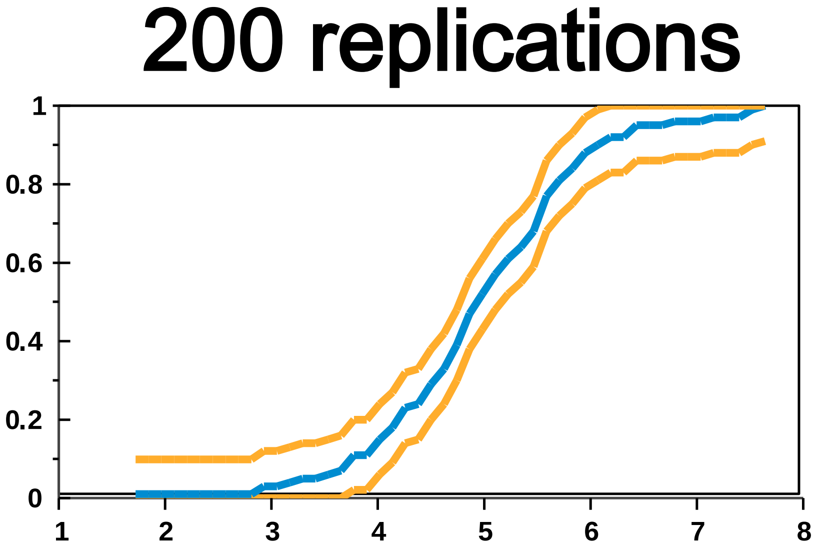
The $\alpha 100\%$ confidence interval for the p^{th} fractile can be estimated by $[Y_i, Y_j]$, where Y_k is the $(n - k + 1)^{th}$ largest value from the Monte Carlo simulation, $i = \text{floor}(np - b)$, $j = \text{ceiling}(np + b)$, $b = z^{-1}((1 - \alpha)/2)\sqrt{(np(1 - p))}$

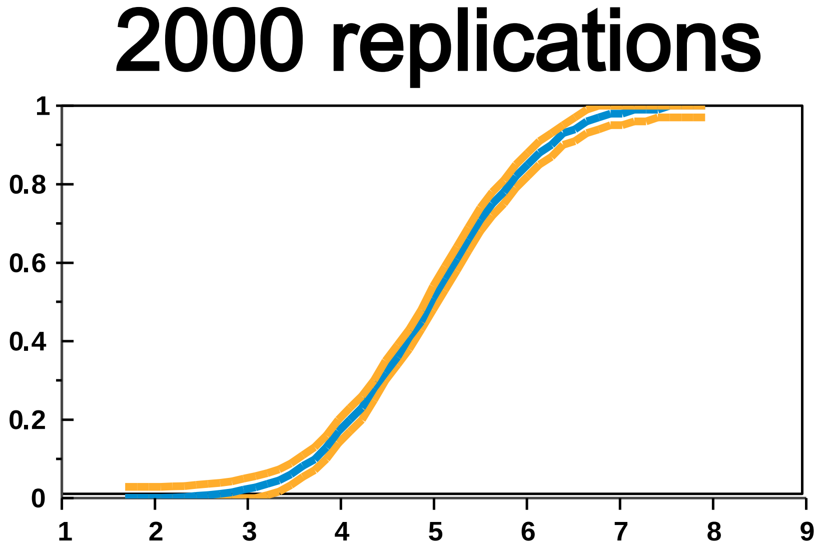
- Vary n to get the precision you desire
- But remember this represents only sampling error, not measurement error

SKIP

Kolmogorov-Smirnov bounds

Bounds on the distribution as a whole





95% of the time, the entire distribution will lie within the bounds

SKIP

Latin hypercube sampling

- Stratify the sampling of distributions
  - Divide each distribution into n equal-probability regions, where n = #reps
  - For each replicate, select one deviate from each distribution
  - Select each region only once
- Often much better for moments, sometimes better for tails
- May reduce the number of replicates needed

Monte Carlo propagation

- Error depends only on replications ($1/\sqrt{N}$)
  - 20% accuracy for 95% of cases (2 sigma) needs $N = 50$
  - 20% accuracy in 99.9% of cases needs $N = 113$
- Doesn't depend on dimensionality of inputs
- Works with arbitrary input distribution shapes
- Can accommodate correlation among inputs
- Works well if uncertainty is *only aleatory* and if you're just interested in the *point value*

Each variable gets a single value

- A variable can't be independent of itself, e.g., a multiple-route exposure model

$$\frac{c_{\mathrm{air}}\, i_{\mathrm{air}}}{BW} + \frac{c_{\mathrm{water}}\, i_{\mathrm{water}}}{BW} + \frac{c_{\mathrm{soil}}\, i_{\mathrm{soil}}}{BW}$$

- All *BW*'s should be the same value within a single Monte Carlo replicate

Multiple instantiation problem

- Multiple instantiation occurs when a Monte Carlo simulation improperly creates several different instances of a single value
- Be wary of “libraries” of partial results
- The multiple instantiation problem can be thought of as a problem in which the modeler failed to realize variables were (perfectly) dependent

Same-plate-of-fish problem

- The opposite problem occurs when a single value is used for what should be many differently valued instances
- For instance, consider a simulation in which distributions for exposure duration, body size and contaminant concentration are simply convolved
  - Implies the *same* concentration at each exposure
  - Also implies you weigh the same at each exposure

Microevent modeling

Monte Carlo simulation

Simulate many anglers

Simulate an angler

Sample *BW*, *ED*

Simulate *ED* years

Simulate a year

Sample *EF*

Simulate *EF* meals

Simulate a meal

Sample C_{fish}, *IR*, *LOSS*

$slug = C_{\text{fish}} \times (1 - LOSS) \times IR \times CF$

$acute = slug \,/\, BW$

$total = total + slug$

$chronic = total \,/\, (BW \times AT)$

Conclusions about Monte Carlo

SKIP

Steps in a Monte Carlo PRA

- Clarify the questions and gather data
- Formulate the model (identify variables and their interrelationships)
- Specify distributions for each input
- Specify dependencies and correlations
- Run simulation
- Conduct sensitivity studies
- Present results
- Discuss limitations of the assessment

SKIP

Pitfalls

- Muddled model
- Poorly chosen distributions, especially tails
- Infeasible correlation matrices
- Inconsistent independence assumptions
- Multiply instantiated variables
- Too few replications
- No sensitivity studies
- Confounding of different populations

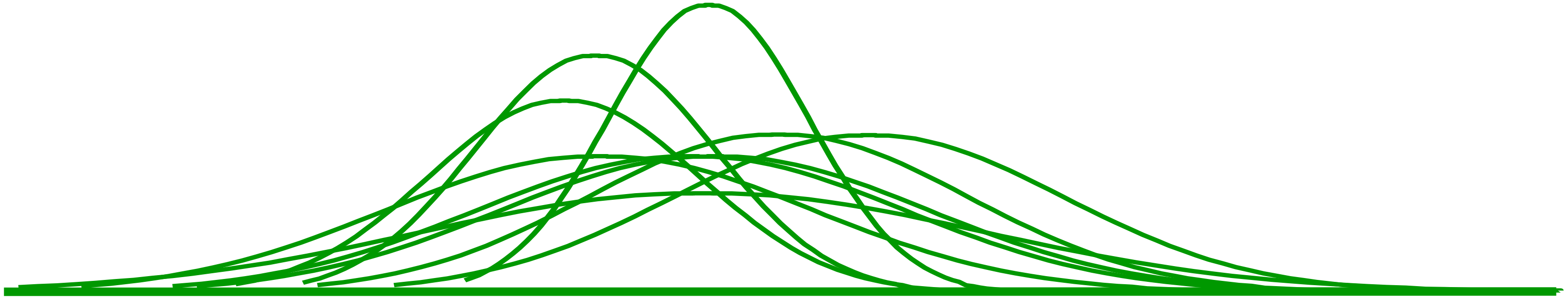
Limitations of Monte Carlo

- Thought by some to be too difficult
- Needs precise, stationary distributions
- Nonlinear dependencies often ignored
- Unknown dependencies almost always ignored
- Assumes model is well understood
- Can be computationally expensive
- Backcalculations clunky
- Confuses incertitude with variability

SKIP

Two-dimensional simulation

- Monte Carlo nested *inside* Monte Carlo
- Inner loop for variability
- Outer loop for uncertainty (incertitude)
- Squared number of replications
- Integrated sensitivity analysis



Monte Carlo in sra.r

- Output
  - <enter variable name>, plot, lines, show, summary
- Characterize
  - mean, sd, var, median, quantile, fivenum, left, right, prob, cut, percentile, iqr, random, range
- Compute
  - exp, log, sqrt, abs, round, trunc, ceiling, floor, sign, sin, cos, tan, asin, acos, atan, atan2, reciprocate, negate, +, -, *, /, pmin, pmax, ^, and, or, not, mixture, smin, smax, complement

Monte Carlo in sra.r

- Construct
  - normal, etc.
  - histogram
  - quantiles
  - MM <tab> <tab>
  - ME <tab> <tab>
  - ML <tab> <tab>

Supported named distributions

bernoulli, beta (B), binomial (Bin), cauchy, chi, chisquared, delta, dirac, discreteuniform, exponential, exponentialpower, extremevalue, F, f, fishersnedecor, fishertippett, fisk, frechet, gamma, gaussian, geometric, generalizedextremevalue (GEV), generalizedpareto, gumbel, histogram, inversegamma, laplace, logistic, loglogistic, lognormal (L), logtriangular, loguniform, negativebinomial, normal (N), pareto, pascal, powerfunction, poisson, quantiles, rayleigh, reciprocal, shiftedloglogistic, skewnormal (SN), student, trapezoidal, triangular (T), uniform (U), weibull

References

Abbas, A.E. 2003. Entropy methods for univariate distributions in decision analysis. *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*: 22nd International Workshop, editied by C.J. Williams, American Institute of Physics.

Cullen, A.C., and H.C. Frey. (1999). *Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. Plenum Press: New York.

Evans, M., N. Hastings, B. Peacock. 2000. *Statistical Distributions*. John Wiley and Sons, New York.

Ferson, S., R.B. Nelsen, J. Hajagos, D.J. Berleant, J. Zhang, W.T. Tucker, L.R. Ginzburg and W.L. Oberkampf. 2004. *Dependence in probabilistic modeling, Dempster-Shafer theory, and probability bounds analysis*. Sandia National Laboratories, Albuquerque, NM.

Hansen, F. (1997). *Policy for Use of Probabilistic Analysis in Risk Assessment at the U.S. Environmental Protection Agency*. EPA/ORD/NCEA May 15:1-4. Available on-line at http://www.epa.gov/osp/spc/probpol.htm.

Jaynes, E.T. (edited by G. Larry Bretthorst). 2003. *Probability Theory: The Logic of Science*. Cambridge University Press.

Lee, R.C. and W.E. Wright. 1994. Development of human exposure-factor distributions using maximum-entropy inference. *Journal of Exposure Analysis and Environmental Epidemiology* 4:329-341.

Morgan, M.G. and M. Henrion. (1990). *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, Cambridge, UK.

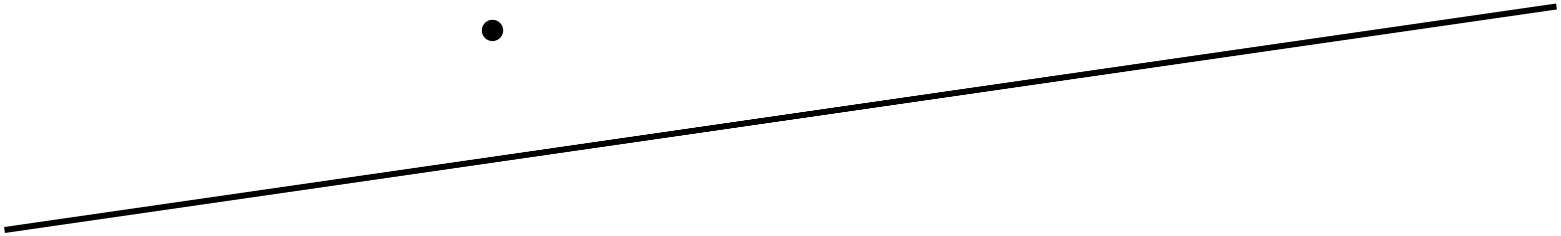
Scheuer, E.M. and D.S. Stoller. (1962). On the generation of normal random vectors. *Technometrics* 4:278-281.

Solana, V. and N.C. Lind. 1990. Two principles for data based on probabilistic system analysis. *Proceedings of ICOSSAR '89, 5th International Conferences on Structural Safety and Reliability*. American Society of Civil Engineers, New York.

Two kinds of uncertainty

Euclid

Given a line in a plane, how many parallel lines can be drawn through a point not on the line?



For over twenty centuries, the answer was *one*

Relax one axiom

- Non-Euclidean geometries say *zero* or *many*
- At first, very controversial
- Mathematics richer and wider applications

Crossroads in uncertainty theory

- There is a kind of uncertainty that cannot be handled by traditional Laplacian probability

- Collision of different views about uncertainty
- Richer math, wider applications
- More reasonable and more reliable results

Two kinds of uncertainty

- Variability
- Aleatory uncertainty
- Type A uncertainty
- Stochasticity
- Randomness
- Chance
- Risk

- Incertitude
- Epistemic uncertainty
- Type B uncertainty
- Ambiguity
- Ignorance
- Imprecision
- True uncertainty

Variability = aleatory uncertainty

- Arises from natural stochasticity
- Variability arises from
  - spatial variation
  - temporal fluctuations
  - manufacturing or genetic differences
- Not reducible by empirical effort

Incertitude = epistemic uncertainty

- Arises from incomplete knowledge

- Incertitude arises from
  - limited sample size
  - mensurational limits (‘measurement uncertainty’)
  - use of surrogate data

- Reducible with empirical effort

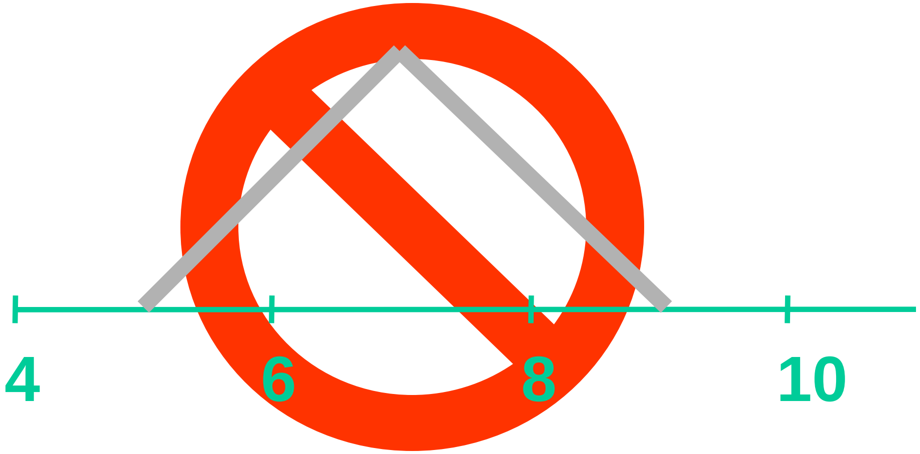
Propagating incertitude

Suppose

A is in [2, 4]

B is in [3, 5]

What can be said about the sum *A*+*B*?



The right answer for risk analysis is [5,9]

Must be treated *differently*

- Variability should be modeled as randomness with the methods of probability theory

- Incertitude should be modeled as ignorance with methods of interval or constraint analysis

- Probability bounding can do both at once

Probability bounds analysis

Bounding probability is an old idea

- Boole and de Morgan
- Chebyshev and Markov
- Borel and Fréchet
- Kolmogorov and Keynes
- Berger and Walley

Why bounding is a good idea

- Often sufficient to specify a decision
- Possible even when estimates are impossible
- Usually easy to compute and simple to combine
- Rigorous, rather than an approximation
- Bounding works with even the crappiest data

(after N.C. Rowe 1988)

Rigorousness

- The computations can be guaranteed to enclose the true results (if the inputs do)
- “Automatically verified calculations”
- You can still be wrong, but the method won’t be the reason if you are

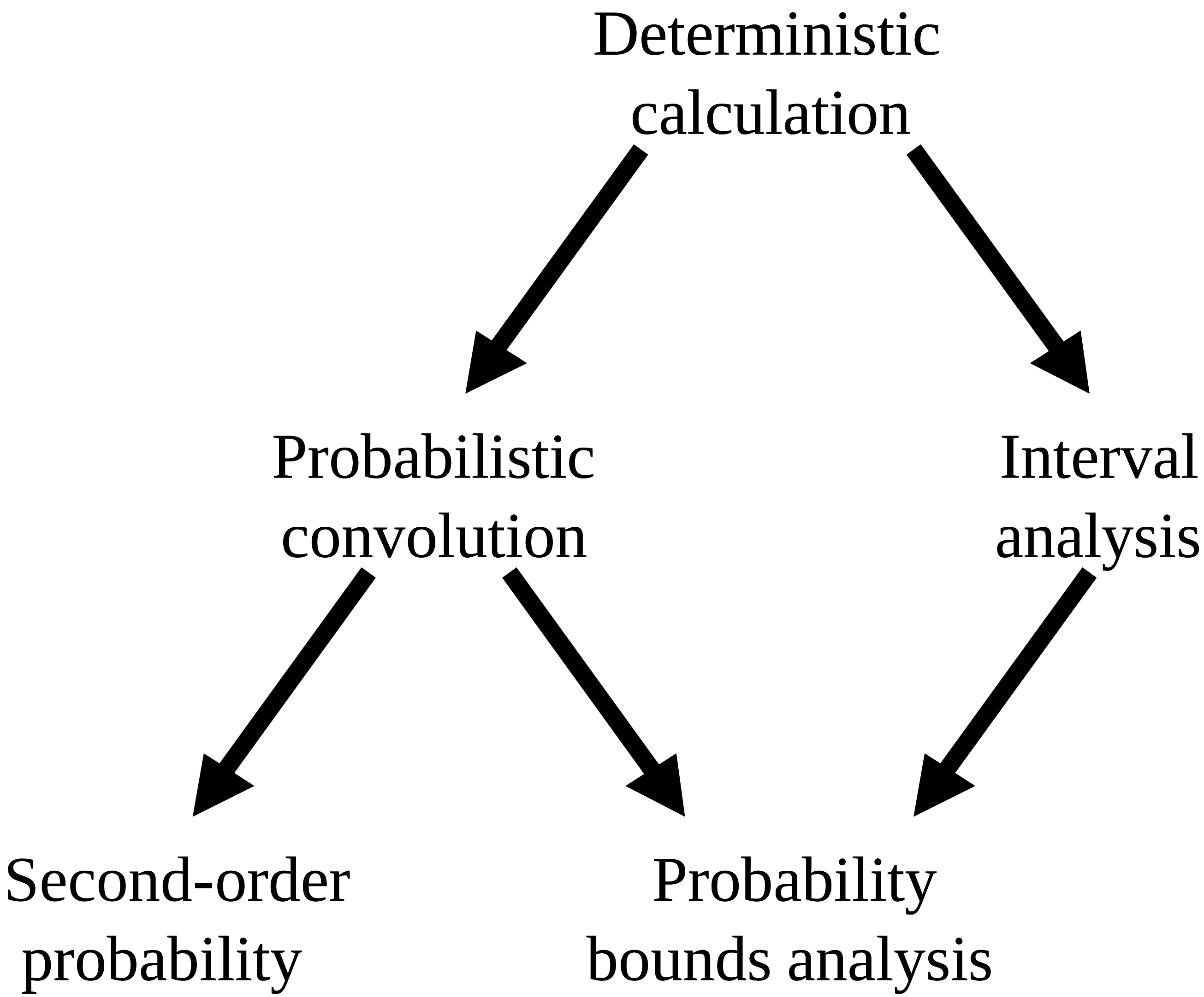
SKIP

Closely related to other ideas

- ***Second-order probability***
  - PBA is easier to work with and more comprehensive
- ***Imprecise probabilities***
  - PBA is somewhat cruder, but a lot easier
- ***Robust Bayesian analysis***
  - PBA does convolutions rather than updating

Bounding approaches like PBA

SKIP



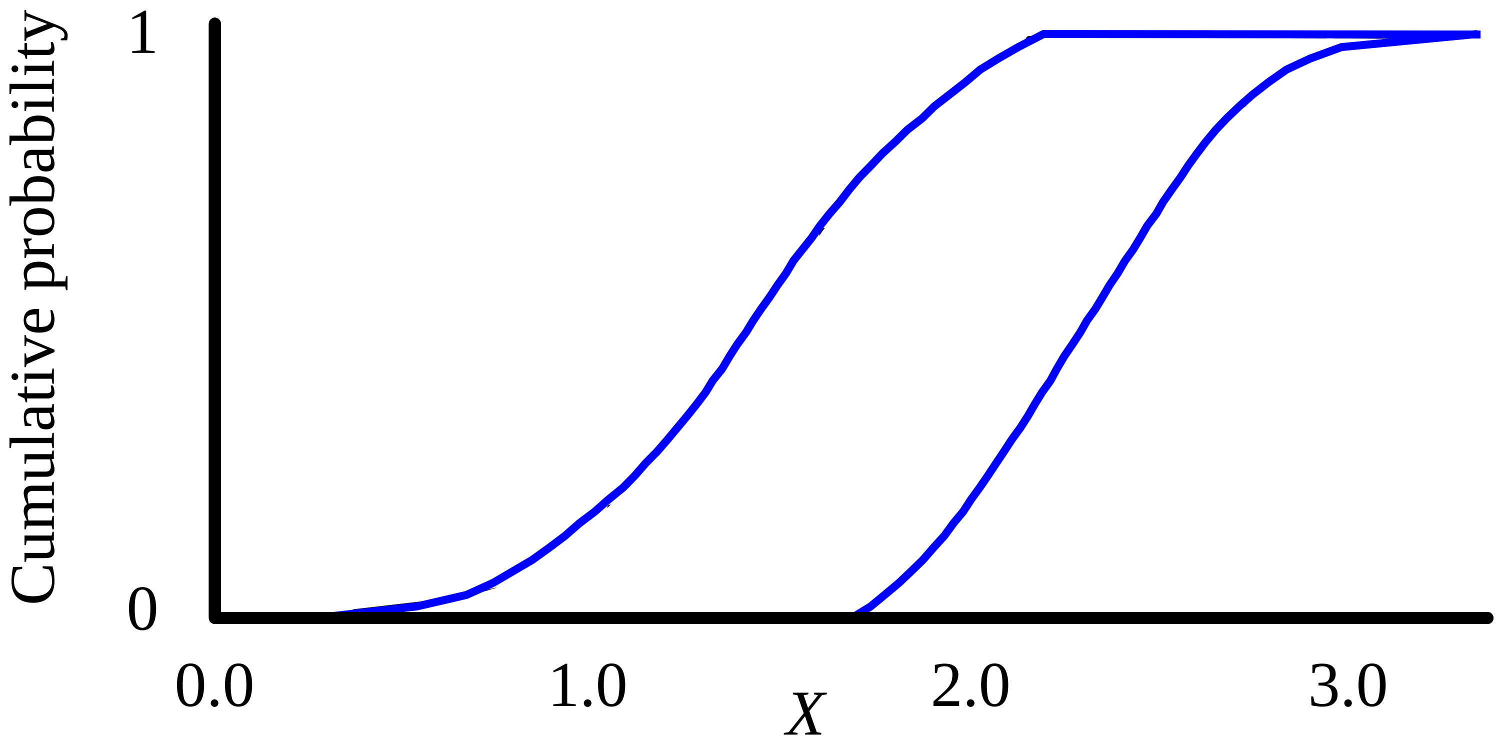
Deterministic calculation
Probabilistic convolution
Interval analysis
Second-order probability
Probability bounds analysis

Probability bounds

- Bridge between qualitative and quantitative
- When data are abundant, it works like probability theory
- When data are sparse, it yields both conservative and optimistic results
- Easy to set up, and cheap to implement

What is a probability box (p-box)?

Interval bounds on a cumulative distribution function (CDF)



Probability bounds analysis

- It’s *not* worst case analysis (distribution tails)
- Marries intervals with probability theory
- Distinguishes variability and incertitude
- Solves many problems in risk analysis
  - Input distributions unknown
  - Imperfectly known correlation and dependency
  - Large measurement error, censoring, small sample sizes
  - Model uncertainty

Calculations

- All standard mathematical operations
  - Arithmetic (+, –, ×, ÷, ^, min, max)
  - Logical operations (and, or, not, if, etc.)
  - Transformations (exp, ln, sin, tan, abs, sqrt, etc.)
  - Backcalculation (deconvolutions, updating)
  - Magnitude comparisons ($<$, $\leq$, $>$, $\geq$, $\subseteq$)
  - Other operations (envelope, mixture, etc.)
- Quicker than Monte Carlo
- Guaranteed to bound the answer
- Optimal solutions often easy to compute

Example: uncontrolled fire

$$F = A \ \& \ B \ \& \ C \ \& \ D$$

Probability of ignition source
Probability of abundant fuel presence
Probability fire detection not timely
Probability of suppression system failure

Imperfect information

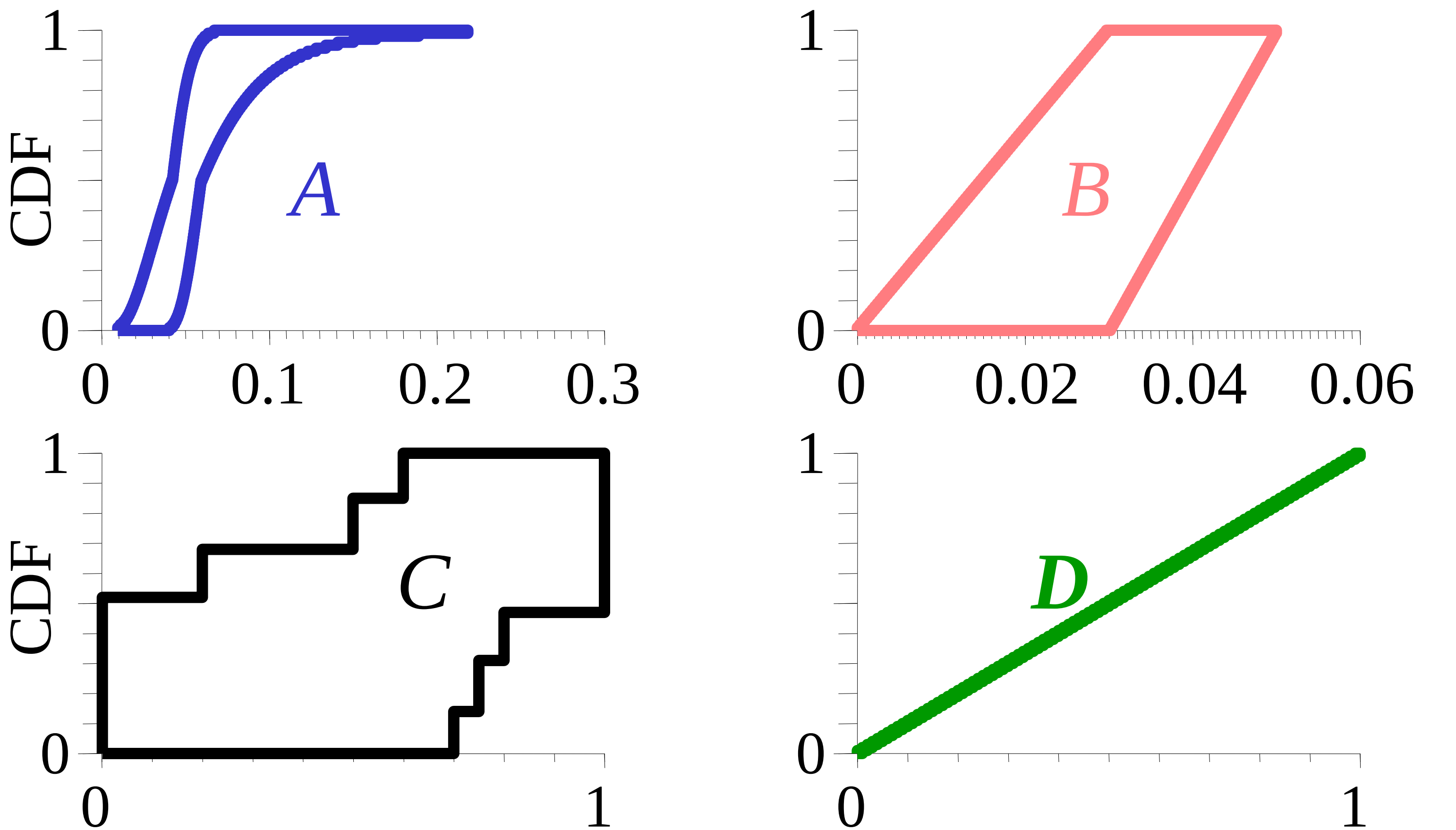
- Calculate $A \,\&\, B \,\&\, C \,\&\, D$, with partial information:
  - A's distribution is known, but not its parameters
  - B's parameters known, but not its shape
  - C has a small empirical data set
  - D is known to be a precise distribution
- Bounds assuming independence?
- Without any assumption about dependence?

A = {lognormal, mean = [.05,.06], variance = [.0001,.001])

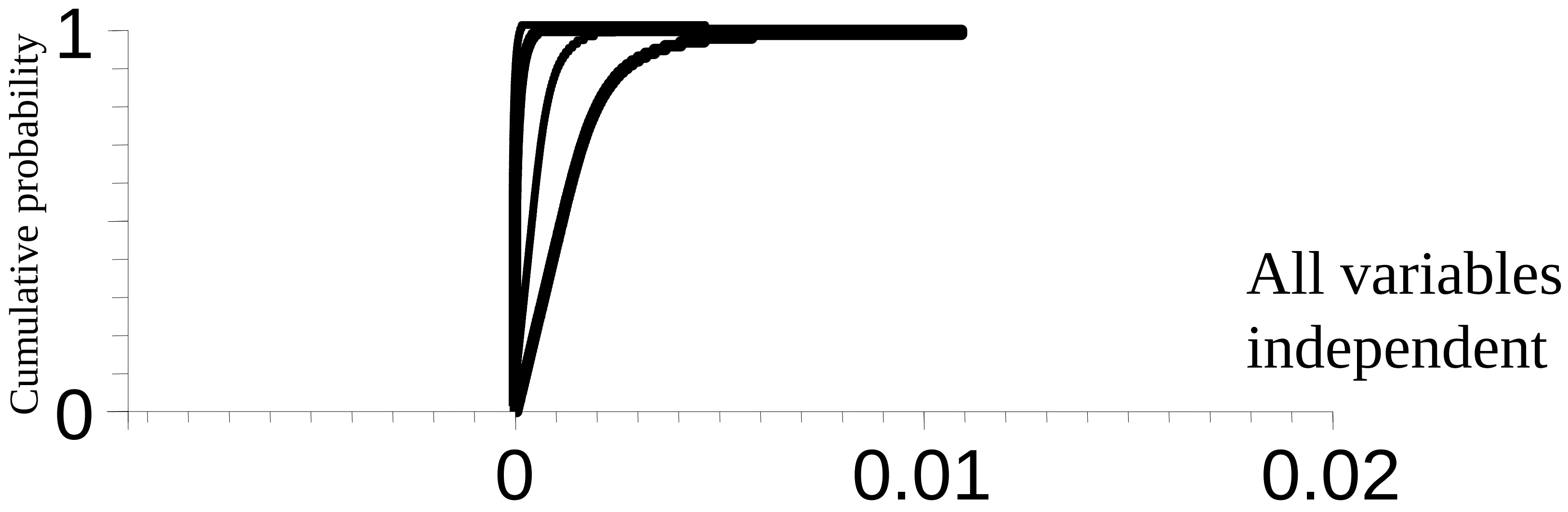
B = {min = 0, max = 0.05, mode = 0.03}

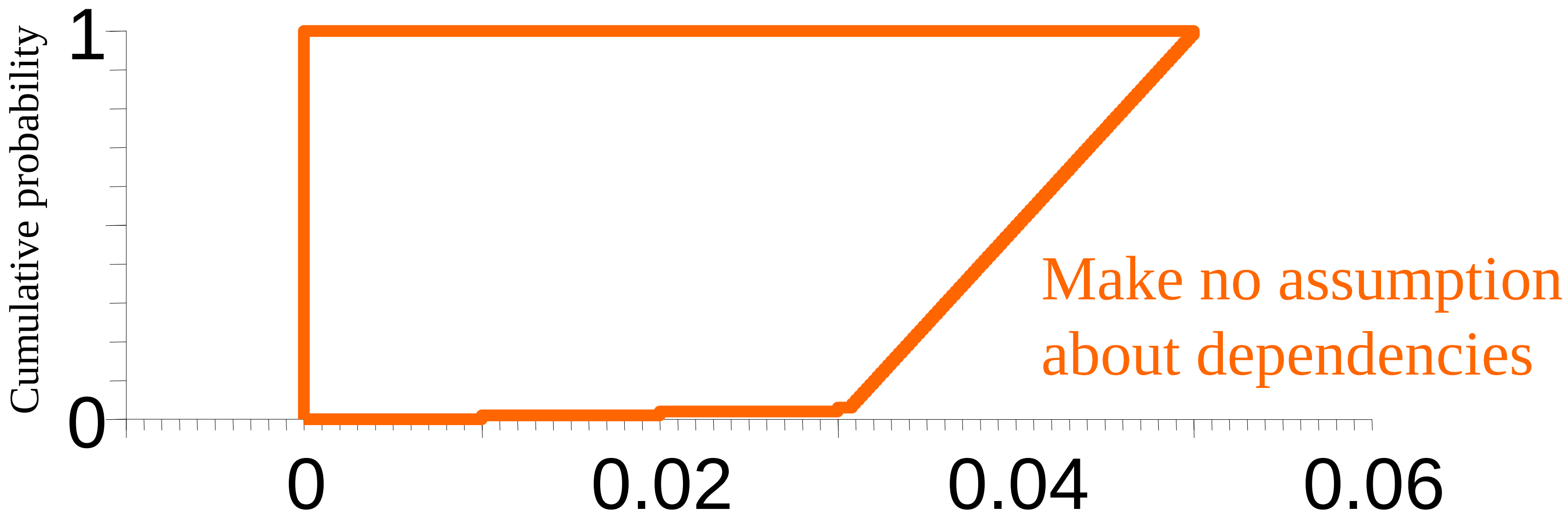
C = {sample data = 0.2, 0.5, 0.6, 0.7, 0.75, 0.8}

D = uniform(0, 1)



Resulting answers





Summary statistics

Independent	
Range	[0, 0.011]
Median	[0, 0.00113]
Mean	[0.00006, 0.00119]
Variance	[2.9×10^{-9}, 2.1×10^{-6}]
Standard deviation	[0.000054, 0.0014]

No assumptions about dependence	
Range	[0, 0.05]
Median	[0, 0.04]
Mean	[0, 0.04]
Variance	[0, 0.00052]
Standard deviation	[0, 0.023]

How to use the results

When uncertainty makes no difference

(because results are so clear), bounding gives confidence in the reliability of the decision

When uncertainty swamps the decision

(*i*) use other criteria within probability bounds, or

(*ii*) use results to identify inputs to study better

Can uncertainty swamp the answer?

- Sure, if uncertainty is huge
- *This should happen* (it’s not “unhelpful”)
- If you think the bounds are too wide, then put in whatever information is missing
- If there isn’t any such information…
  do you want the results to mislead people?

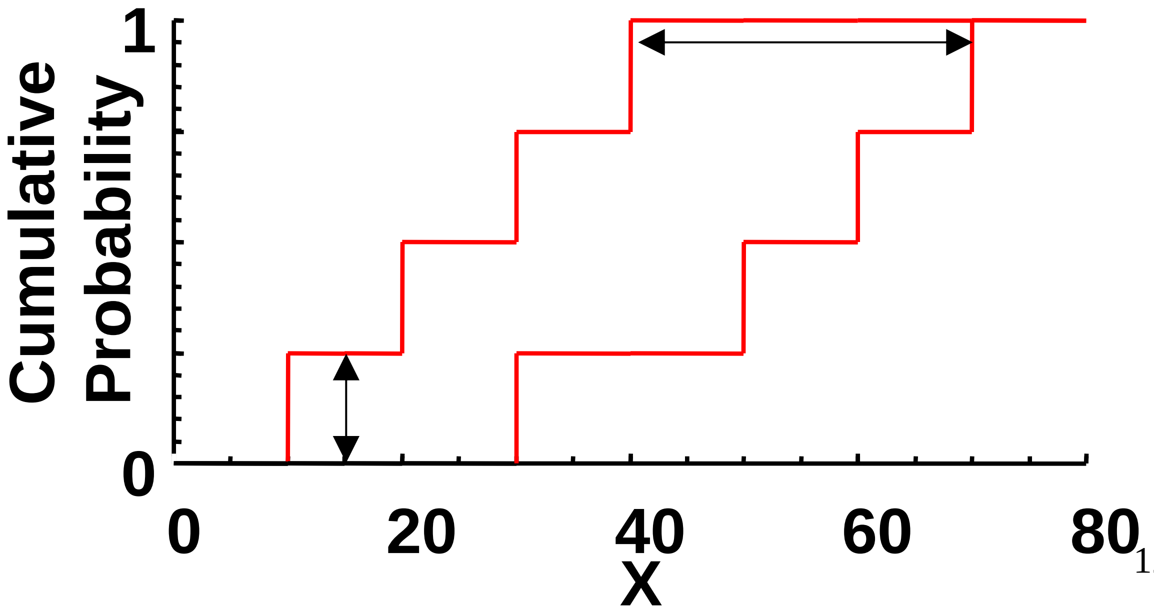
Duality

- Bounds on the probability at a value

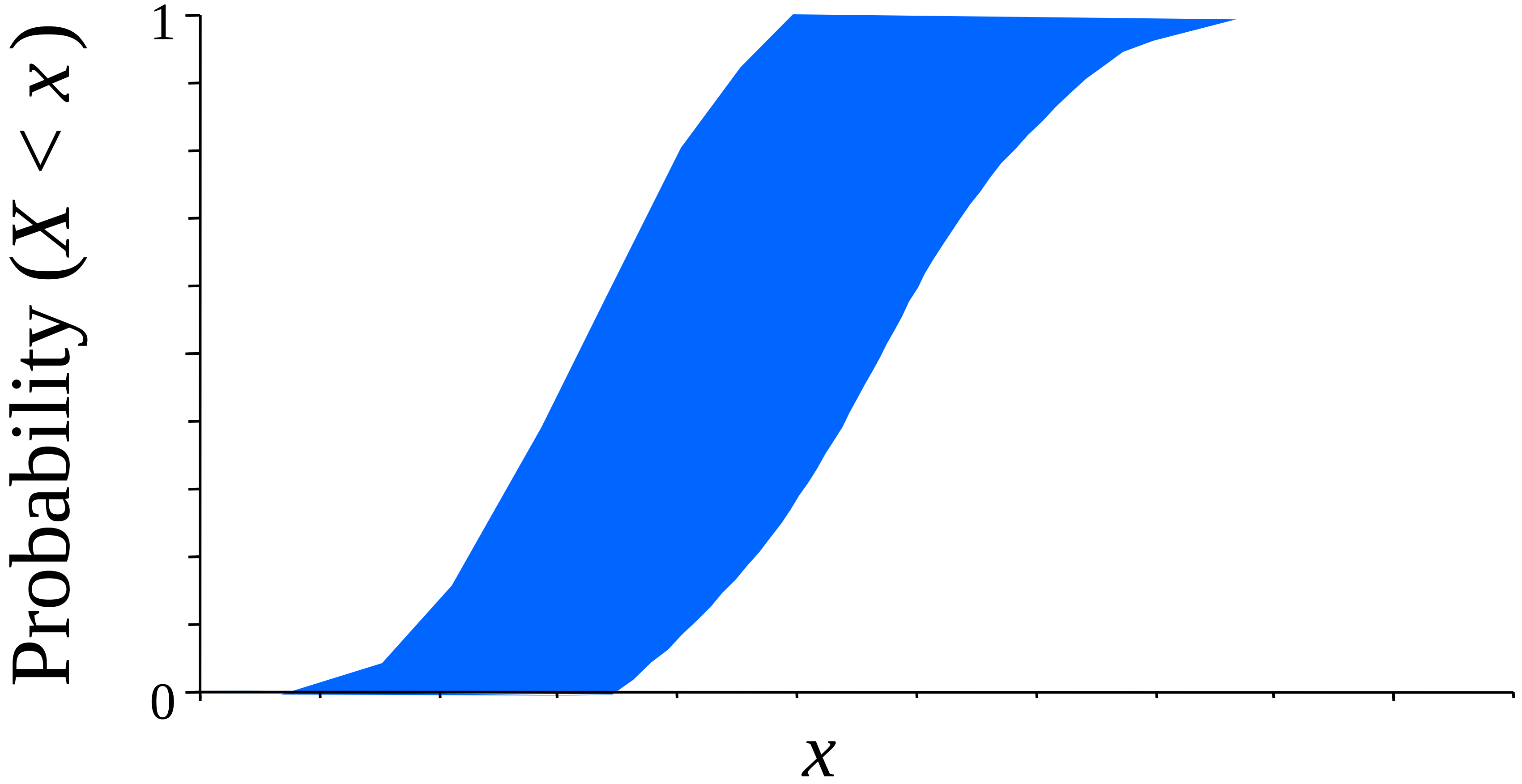
  Chance the value will be 15 or less is between 0 and 25%

- Bounds on the value at a probability

  95th percentile is between 40 and 70

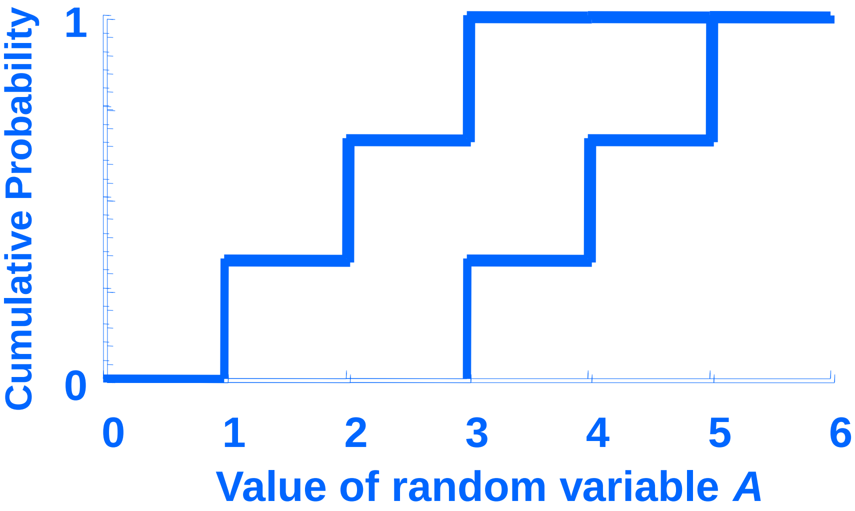


Breadth (incertitude) v. tilt (variability)

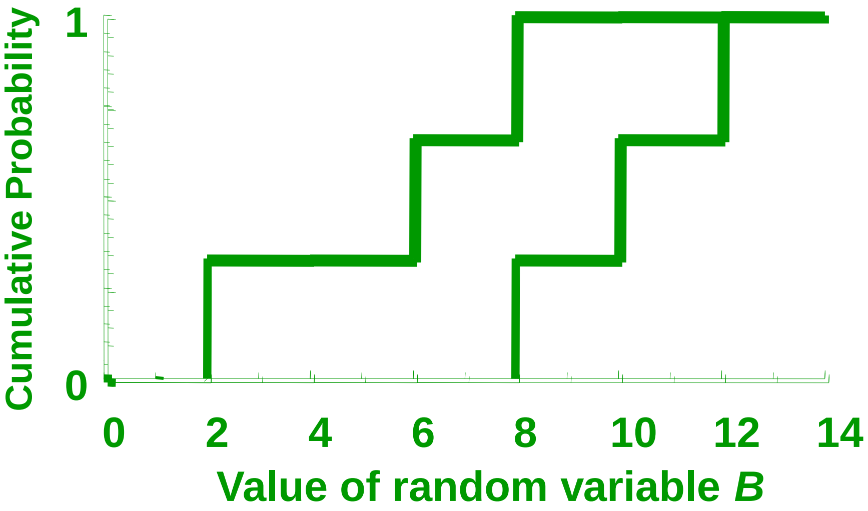


Probability bounds arithmetic

P-box for random variable *A*



P-box for random variable *B*

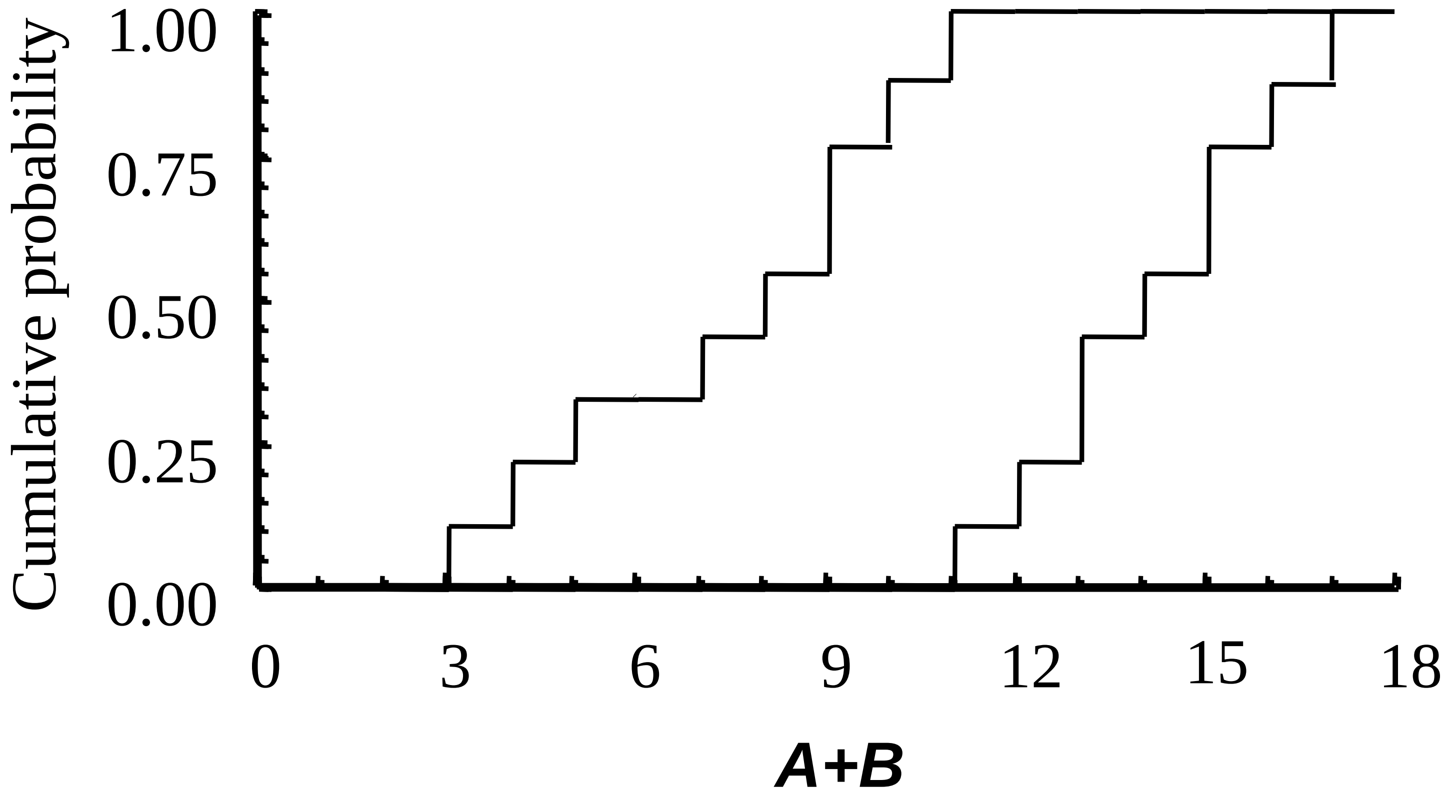


What are the bounds on the distribution of the sum of *A*+*B*?

Cartesian product

$A+B$ independence	$A\in[1,3]$ $p_1 = 1/3$	$A\in[2,4]$ $p_2 = 1/3$	$A\in[3,5]$ $p_3 = 1/3$
$B\in[2,8]$ $q_1 = 1/3$	$A+B\in[3,11]$ prob=1/9	$A+B\in[4,12]$ prob=1/9	$A+B\in[5,13]$ prob=1/9
$B\in[6,10]$ $q_2 = 1/3$	$A+B\in[7,13]$ prob=1/9	$A+B\in[8,14]$ prob=1/9	$A+B\in[9,15]$ prob=1/9
$B\in[8,12]$ $q_3 = 1/3$	$A+B\in[9,15]$ prob=1/9	$A+B\in[10,16]$ prob=1/9	$A+B\in[11,17]$ prob=1/9

A+B under independence

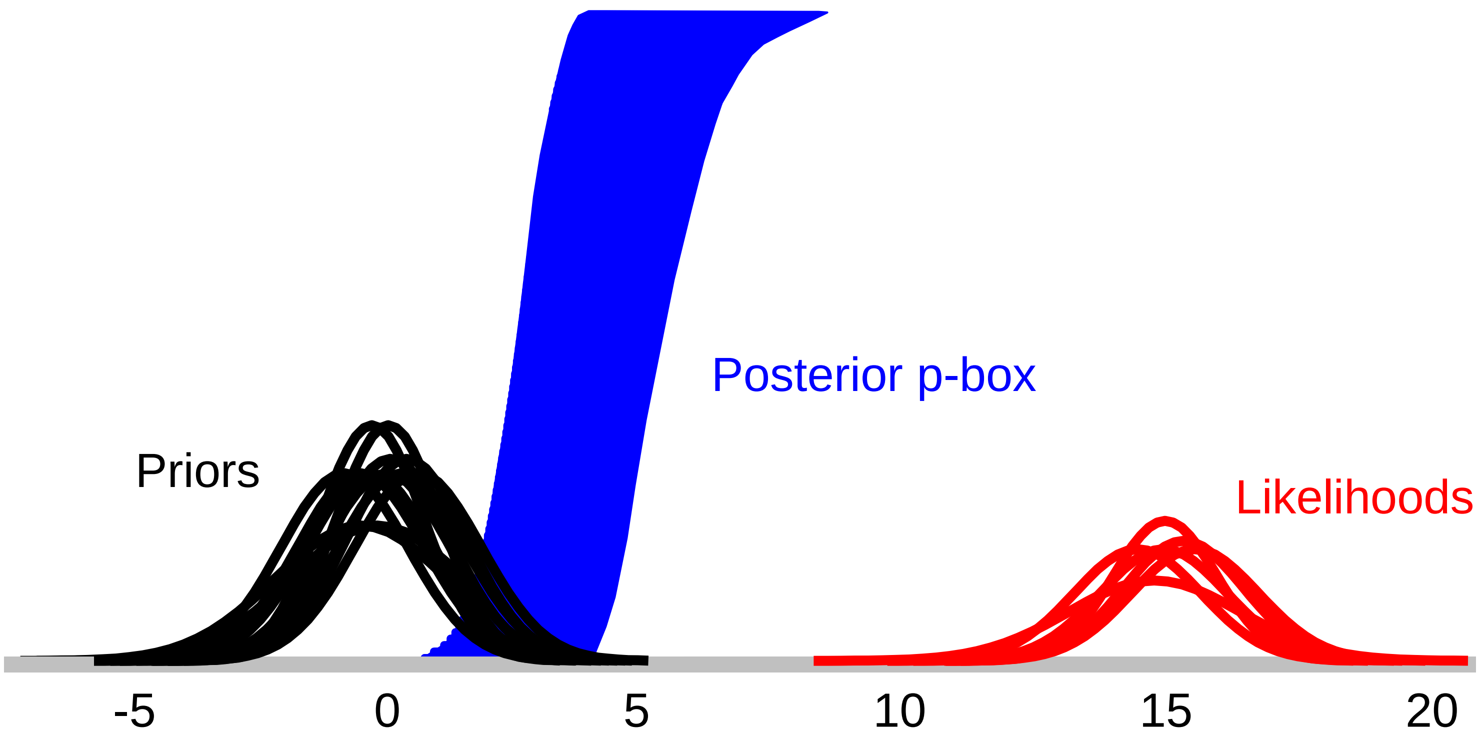


Where do inputs come from?

Where do input p-boxes come from?

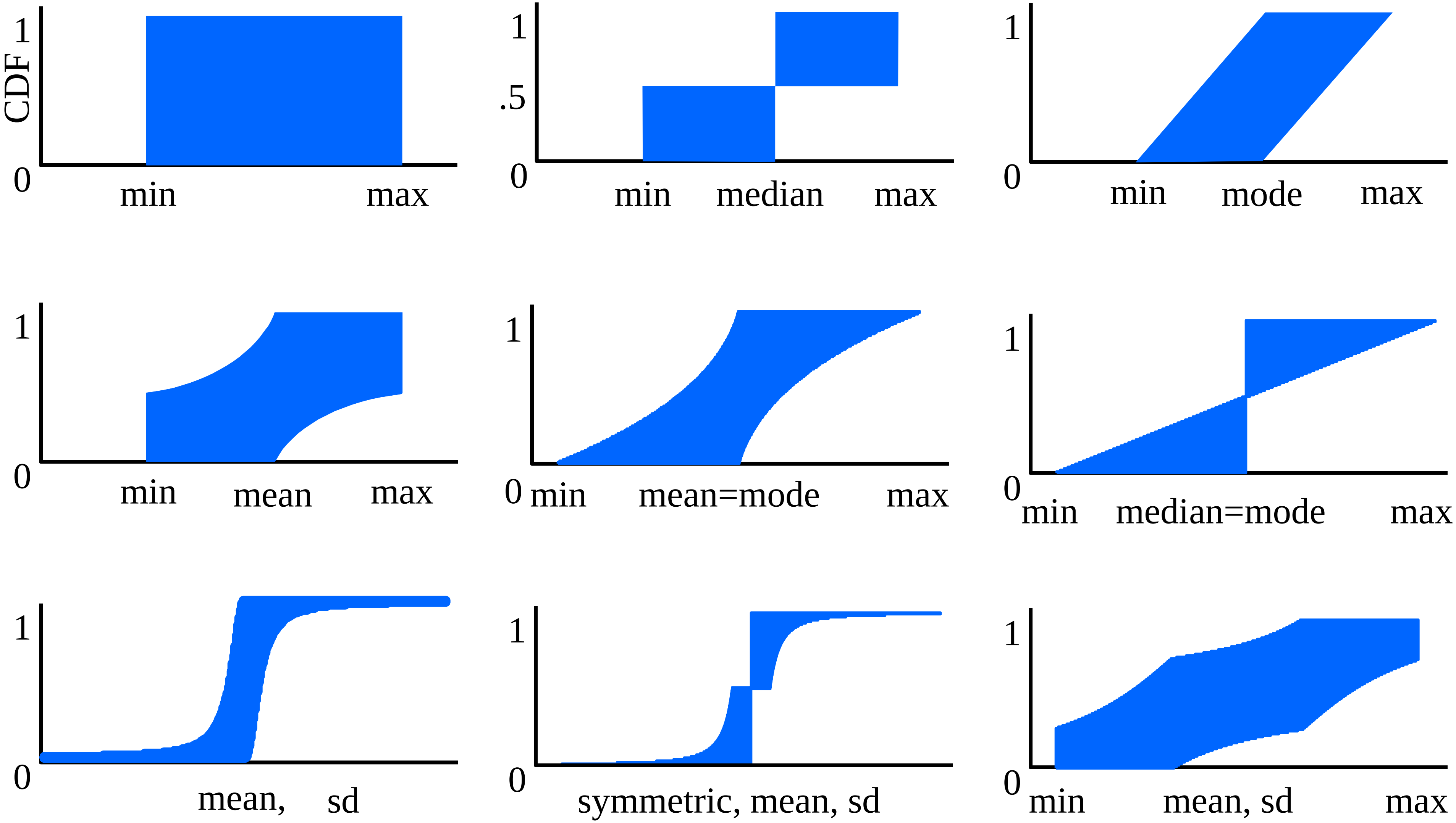
- Prior modeling
  - Uncertainty about dependence
  - Robust Bayes analysis
- Constraint information
  - Summary publications lacking original data
- Sparse or imprecise data
  - Shallow likelihood functions to maximize
  - Measurement uncertainty, censoring, missing data
  - Inferential (sampling) uncertainty

Robust Bayes can make a p-box

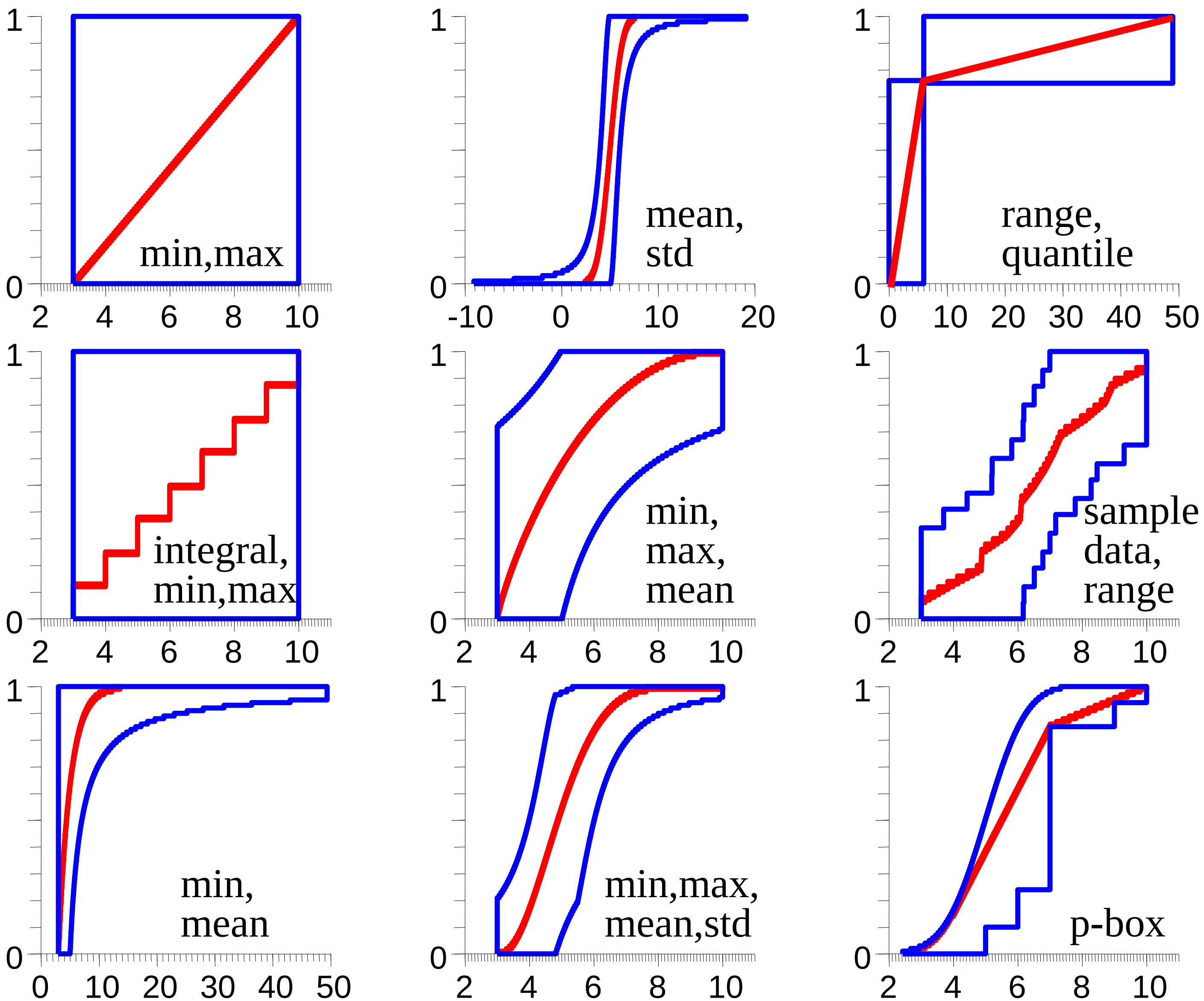


class of priors, class of likelihoods $\Rightarrow$ class of posteriors

Constraint propagation



Comparing p-boxes with maximum entropy distributions

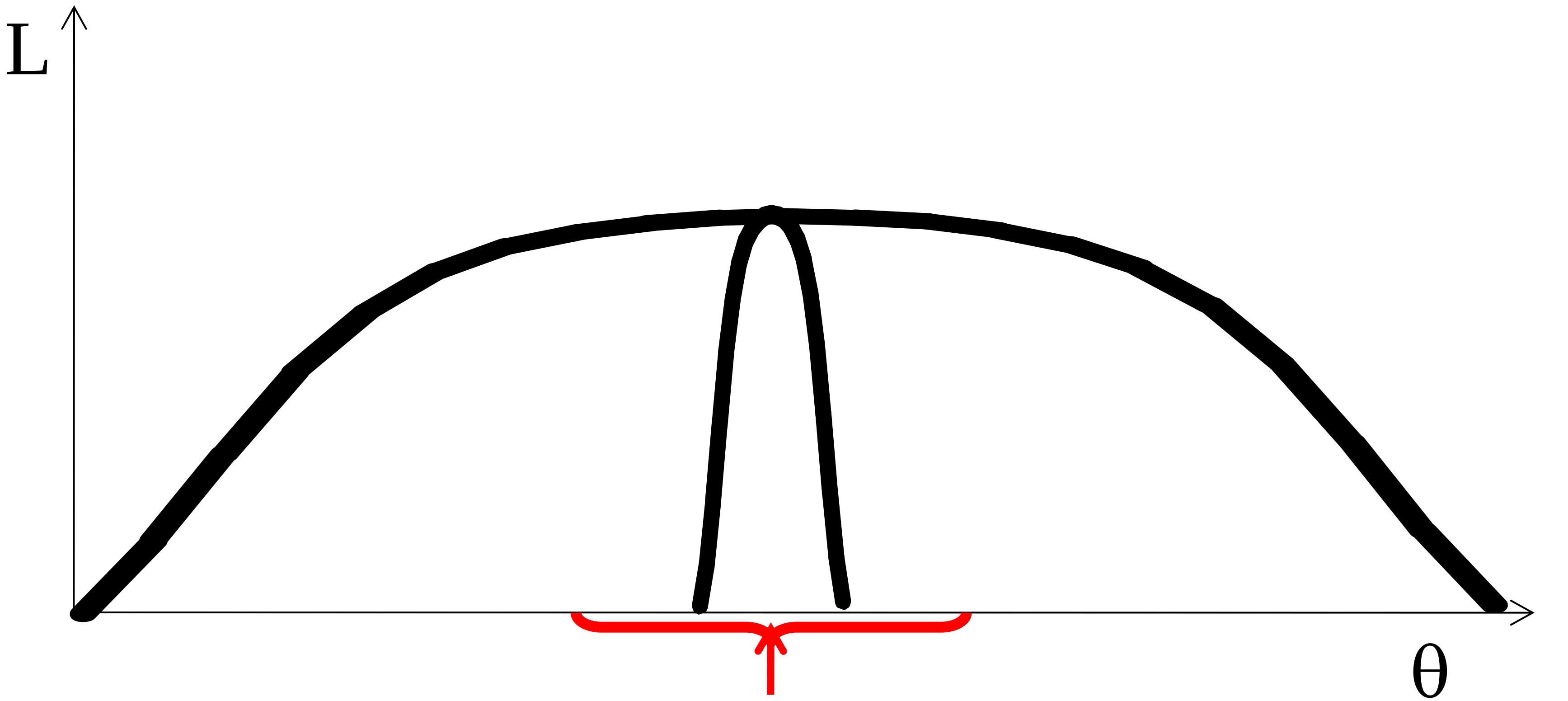


Maximum entropy's problem

- Depends on the choice of scale
- A solution in terms of degradation rate is incompatible with one based on half life even though the information is equivalent
- P-boxes are the *same* whichever scale is used

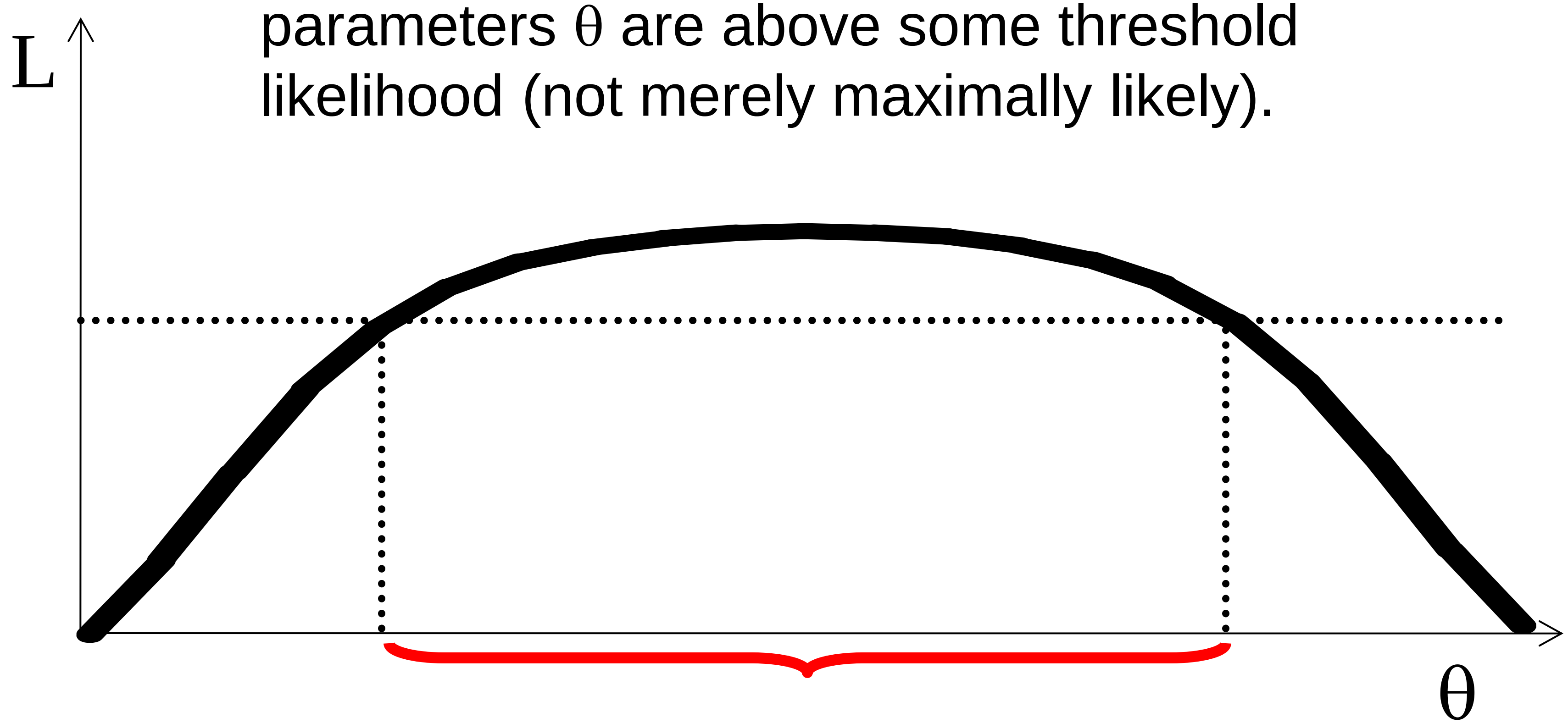
Warner North interprets Ed Jaynes as saying that "two states of information that are judged to be equivalent should lead to the same probability assignments". Maxent doesn't do this! But PBA does.

Sparse data yield shallow likelihoods



Treelining likelihood makes a p-box

The p-box encloses all distributions whose parameters θ are above some threshold likelihood (not merely maximally likely).

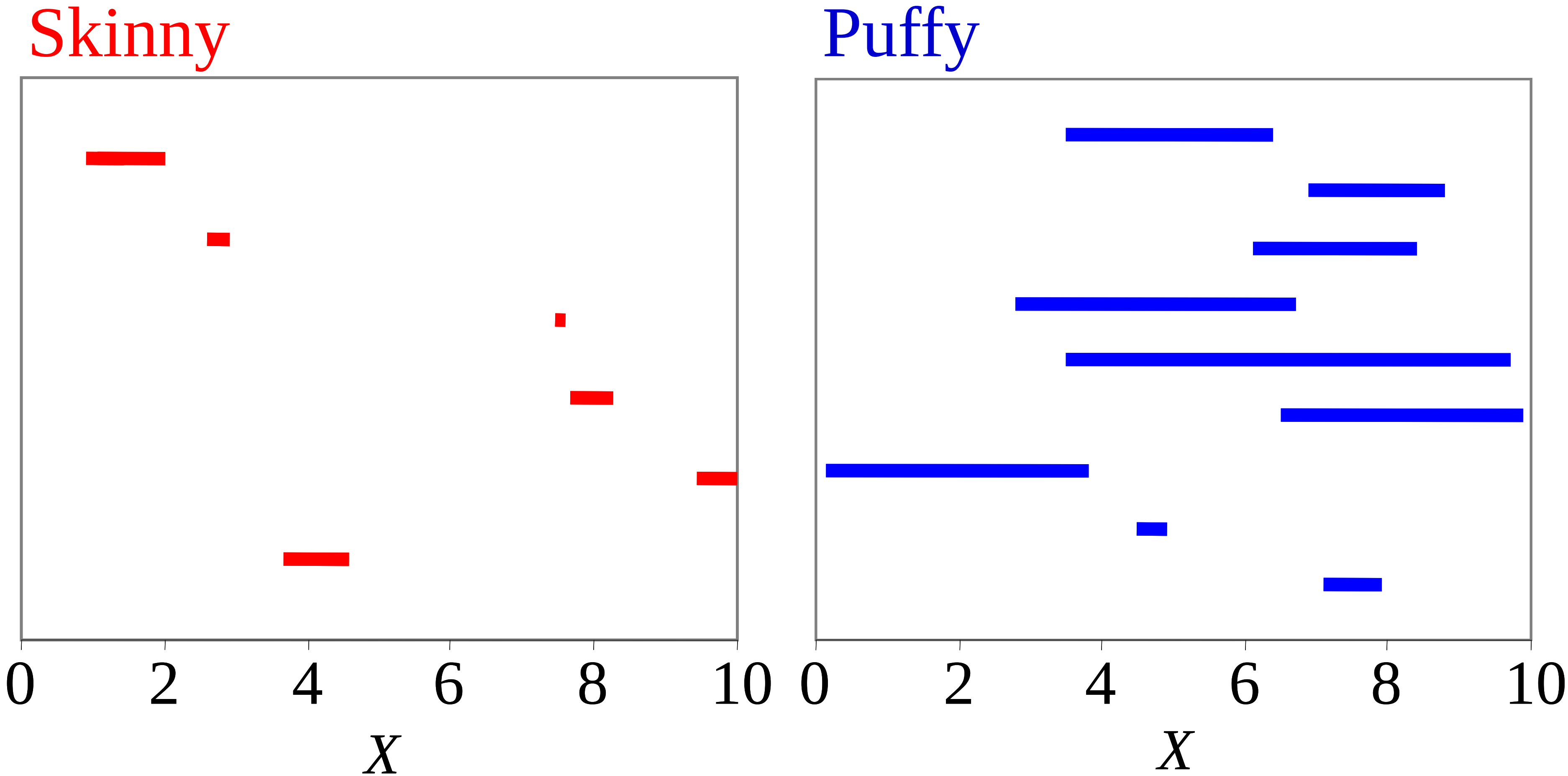


Sources of incertitude in data

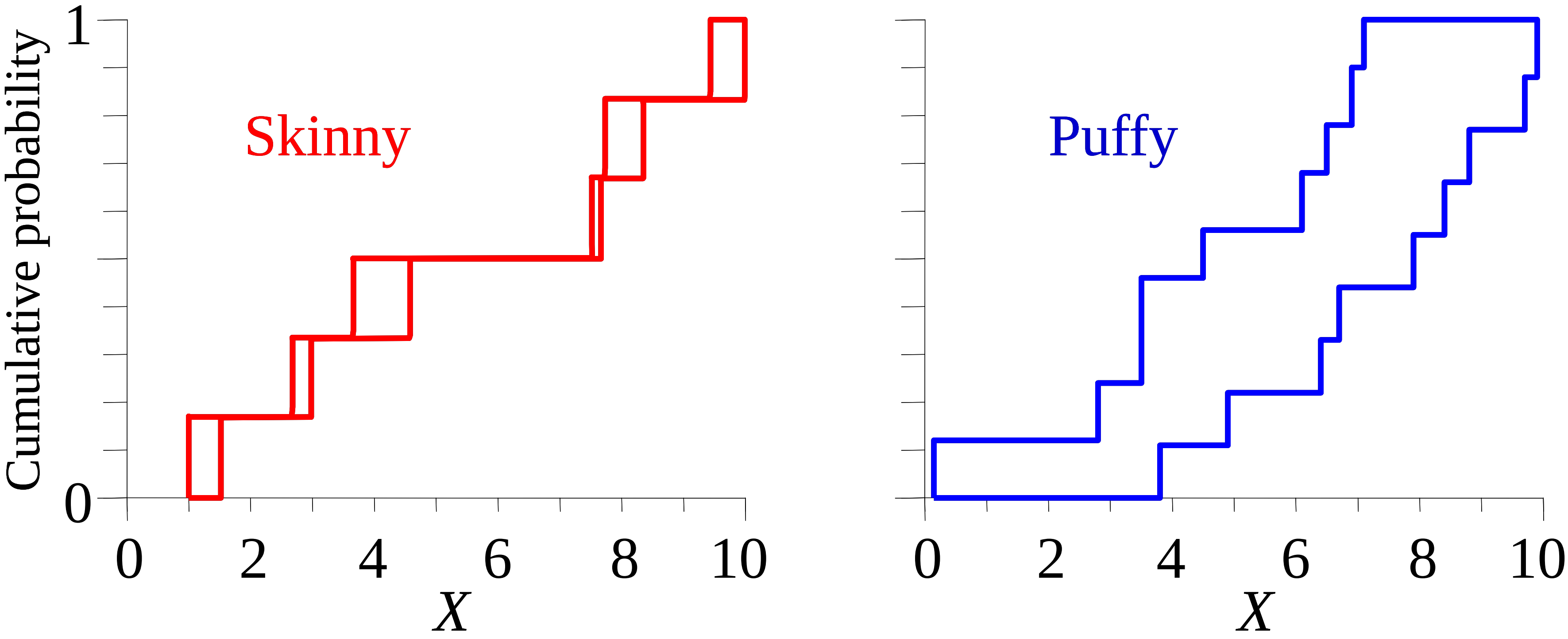
- Periodic observations
  When did the fish in my aquarium die during the night?
- Plus-or-minus measurement uncertainties
  Coarse measurements, measurements from digital readouts
- Non-detects and data censoring
  Chemical detection limits, studies prematurely terminated
- Privacy requirements
  Epidemiological or medical information, census data
- Theoretical constraints
  Concentrations, solubilities, probabilities, survival rates
- Bounding studies
  Presumed or hypothetical limits in what-if calculations

Imprecise sample data

Skinny data	**Puffy data**
[1.00, 2.00]	[3.5, 6.4]
[2.68, 2.98]	[6.9, 8.8]
[7.52, 7.67]	[6.1, 8.4]
[7.73, 8.35]	[2.8, 6.7]
[9.44, 9.99]	[3.5, 9.7]
[3.66, 4.58]	[6.5, 9.9]
	[0.15, 3.8]
	[4.5, 4.9]
	[7.1, 7.9]

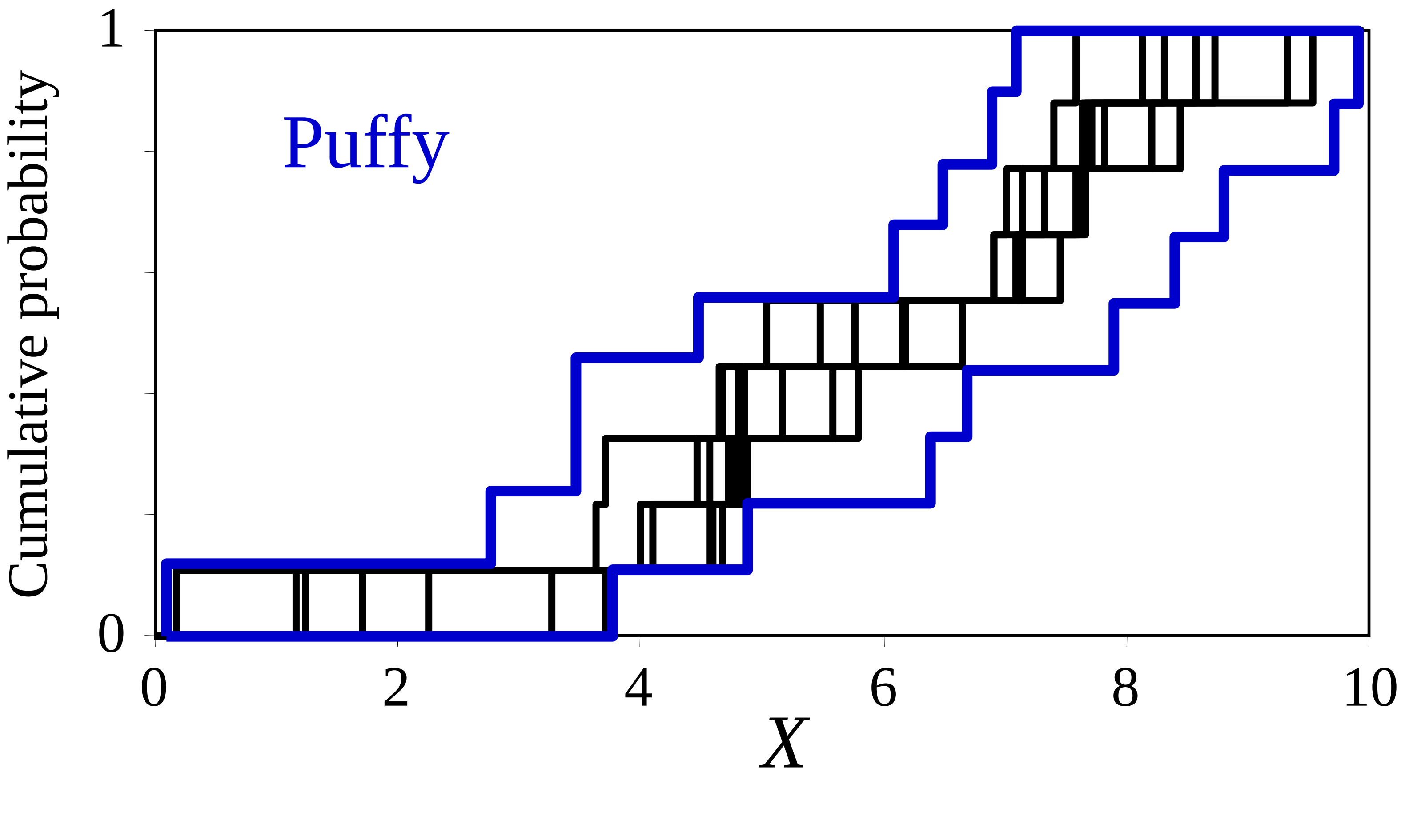

Skinny
Puffy
0
2
4
6
8
10
X
0
2
4
6
8
10
X

Empirical distribution of intervals

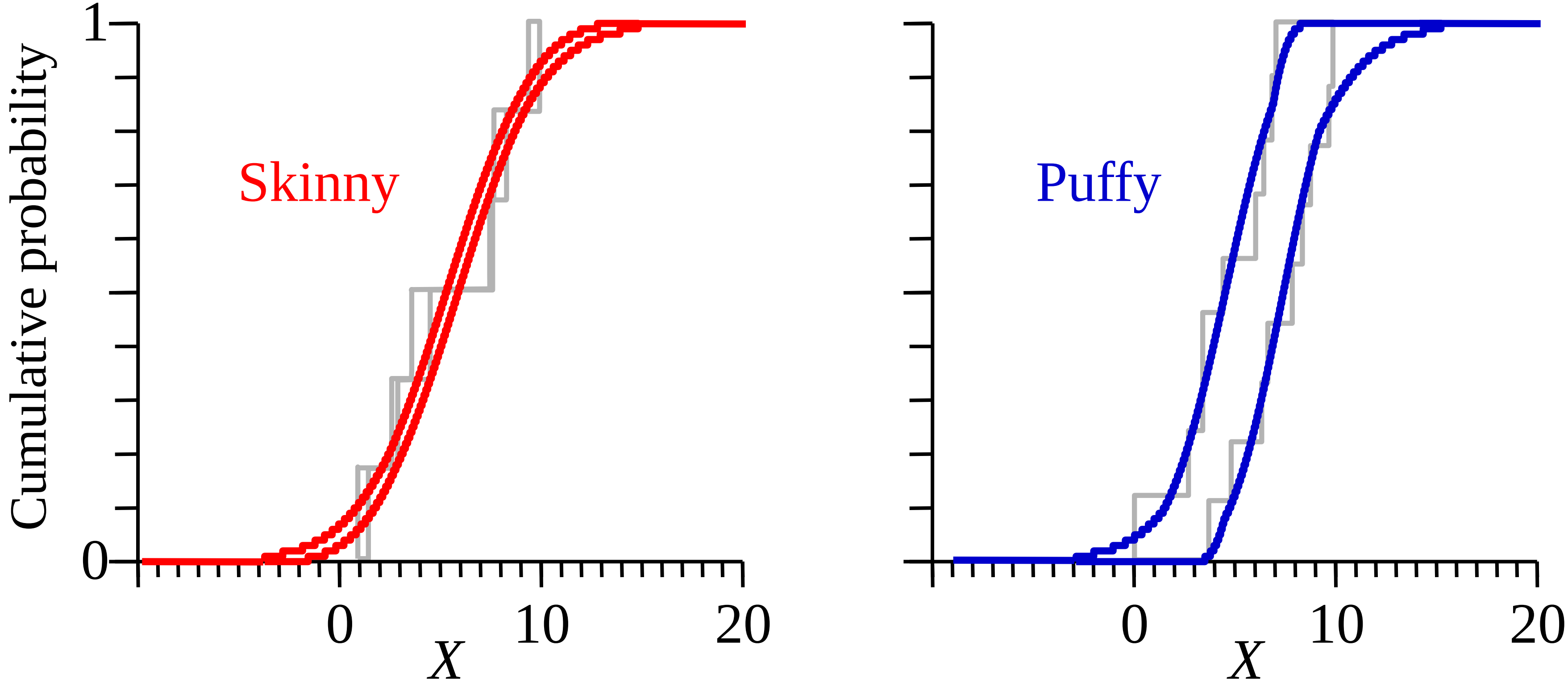


- Each side is cumulation of respective endpoints
- Represents both uncertainty and variability

Uncertainty about the EDF



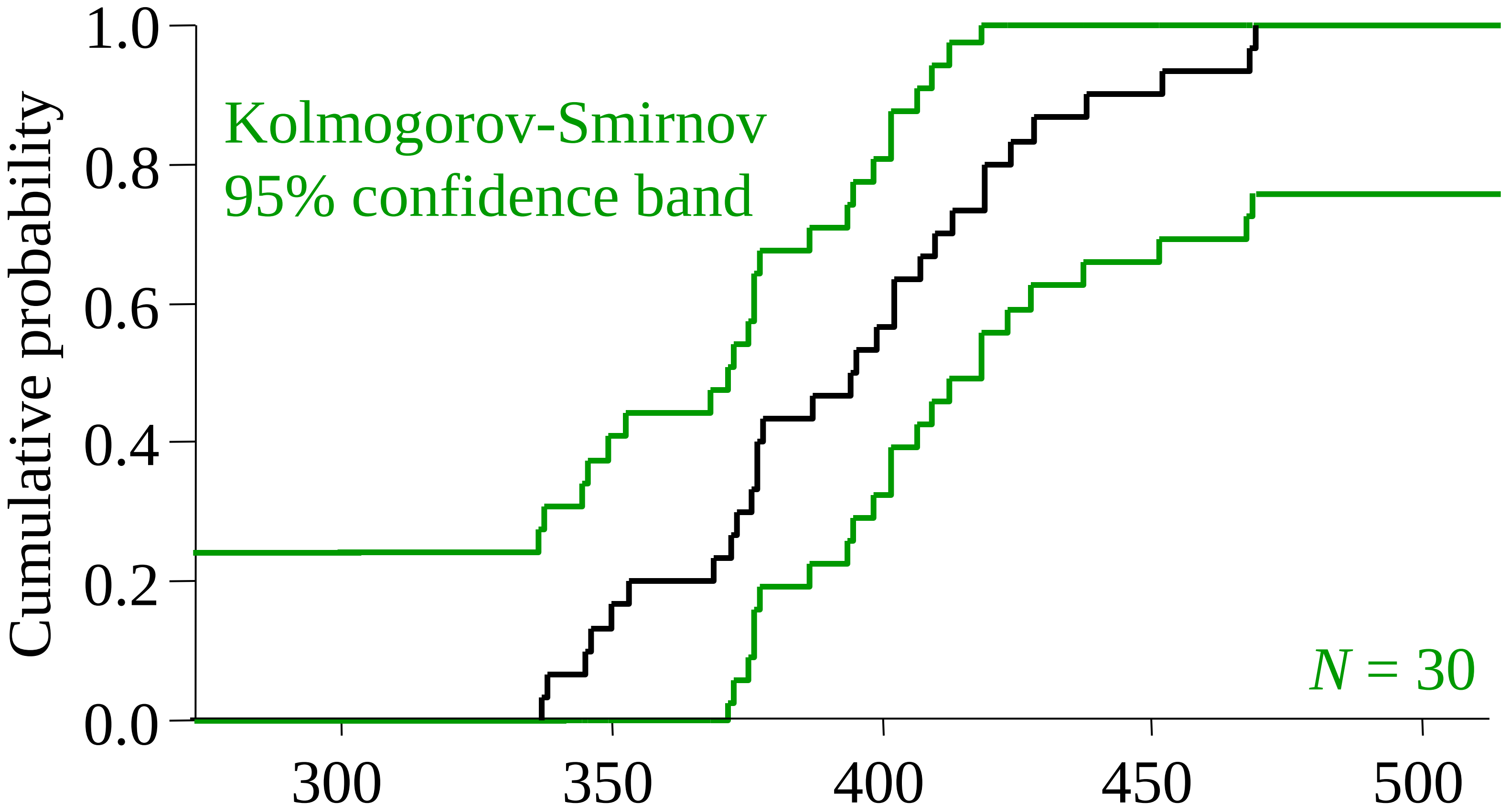
Fitted to normals



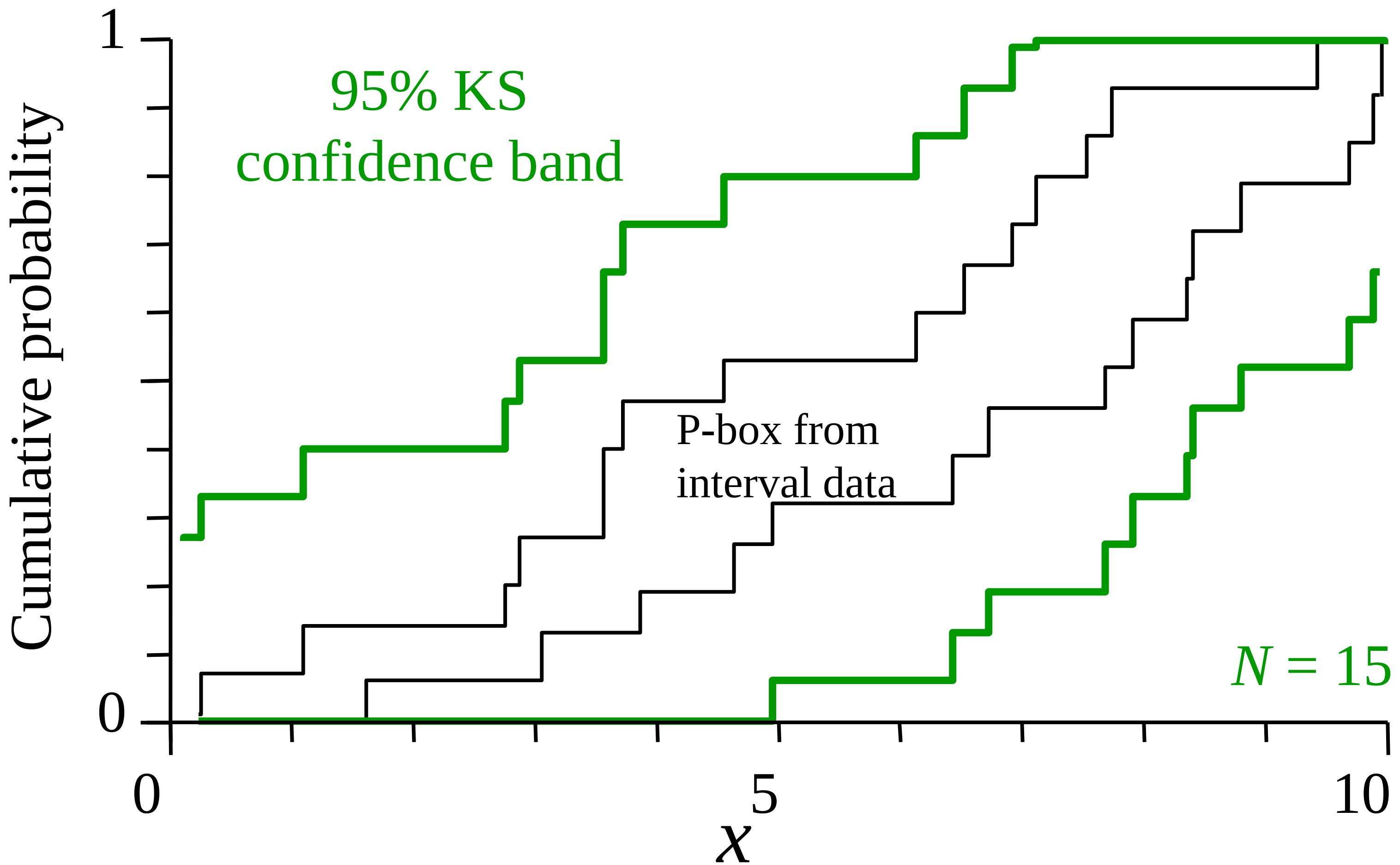
Distributional uncertainty

- Can we trust the raw data or any distribution fitted to them?
- Should account for *sampling uncertainty* about a probability distribution given sampled values
- Kolmogorov-Smirnov confidence procedure
  – 95% (or whatever) confidence for distribution
  – Assumes continuous distribution
  – Random samples, which come from random inputs

Distributional confidence limits



Confidence band about a p-box



SKIP

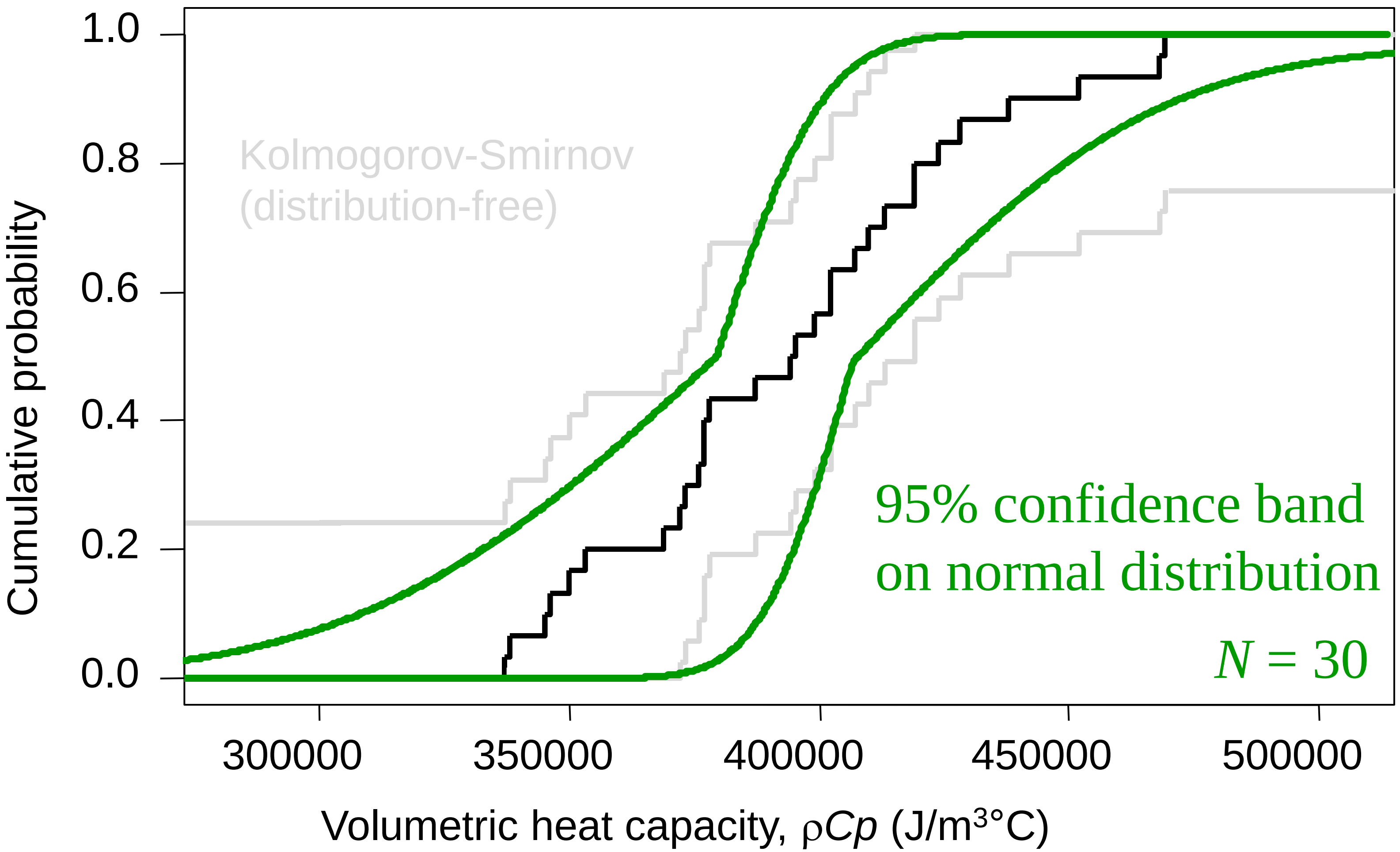
KS generalizes well

- Works with arbitrary input distribution shapes
- Can handle incertitude too
- Generalizes to multivariate outputs
- Conclusion is distribution-free
- Assuming output shape could tighten bounds

SKIP

Distributional confidence limits

(Chen and Iles; Basu)



SKIP

Single-sample confidence band

- Rodríguez described confidence intervals for mean and standard deviation assuming normality from *only one random sample*

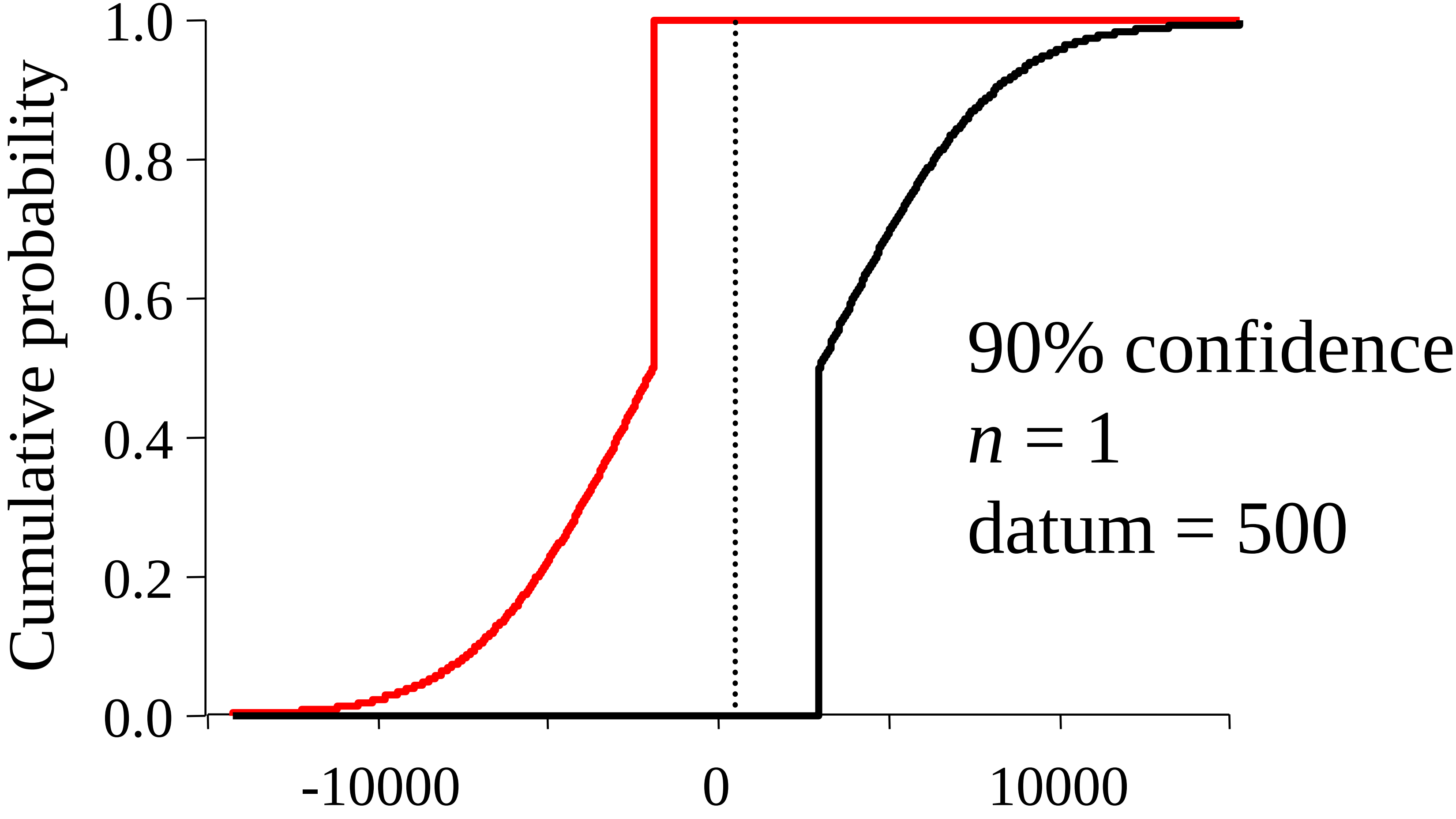
- Combine them to get confidence band

```
onepointconfidenceband.normal = function(x, c=0.95) {
  stopifnot(length(x)==1)
  tm = c(4.83952, 9.678851, 48.39413)
  ts = c(8,17,70)
  k = which(c==c(0.9, 0.95, 0.99))
  normal(x+tm[[k]]*abs(x)*interval(-1,1),interval(0,ts[[k]]*abs(x)))
  }
```

creates a true confidence band

SKIP

Very wide, even assuming normality



Caveat

SKIP

Confidence bands aren't rigorous

- Not compatible with interval arithmetic
  - You can't intersect or compute with them
  - Not rigorous…only statistical
- Could artfully use the Fréchet inequality
  - Combining two bands at 95% yields a 90% band
  - Two 97.5% → 95%, three 95% → 85%, three 98.35% → 95%
- Could *assume* the confidence band is rigorous
  - Assumption often used (e.g., EPA uses a UCL as the EPC)

SKIP

Every p-box represents assumptions

- Constraint p-boxes assumed you knew those parameters
- Nothing intrinsically different about the assumption that converts a confidence band to a p-box

Confidence boxes

- Structures that let you infer confidence intervals (at any confidence level) for a parameter
- Extend confidence distributions (like Student's *t*)
- Different from confidence bands
- Can be propagated as ordinary probability bounds

Example: binomial rate

- Inference about a probability from binary data, k successes out of n trials

$$p \sim \mathrm{env}(\mathrm{beta}(k,\, n-k+1),\, \mathrm{beta}(k+1, n-k))$$

- Similar to robust Bayes analysis but does not require a prior or family of priors

Binomial rate *p* for *k* of *n* trials

$$p \sim \text{env}(\text{beta}(k,\ n-k+1),\ \text{beta}(k+1, n-k))$$

Example

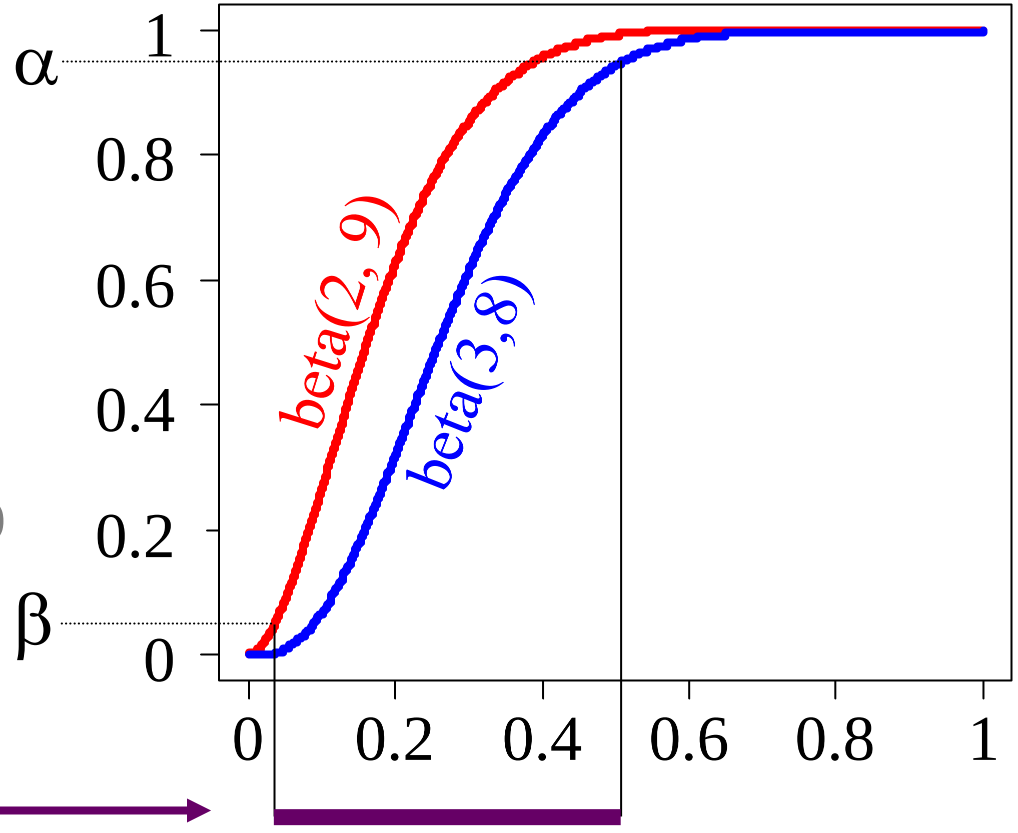
$k = 2$

$n = 10$

CBbinomial(2,10)

CBbernoulli(c(0,0,1,0,1,0,0,0,0,0))

(α–β)100% confidence interval for *p*

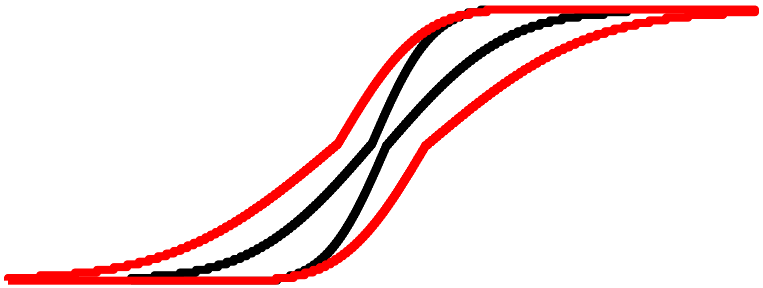


Uncertainties expressible with p-boxes

- Sampling uncertainty (from small *N*)
  - Distributional confidence bands, confidence boxes
- Measurement incertitude
  - Plus-minus ranges, censoring intervals
- Uncertainty about distribution shape
  - Constraints (non-parametric p-boxes)
- Surrogacy uncertainty (have *X* but want *Y*)
  - Modeling

Modeling for surrogacy

- Sometimes you *correct* the best estimate
  - Subtract the weight of the dish when they charge for salad in the cafeteria (+)
  - Sulfur hexafluoride is worth so much CO_2 (×)
- If you're not sure exactly how, then you should *widen the uncertainty*

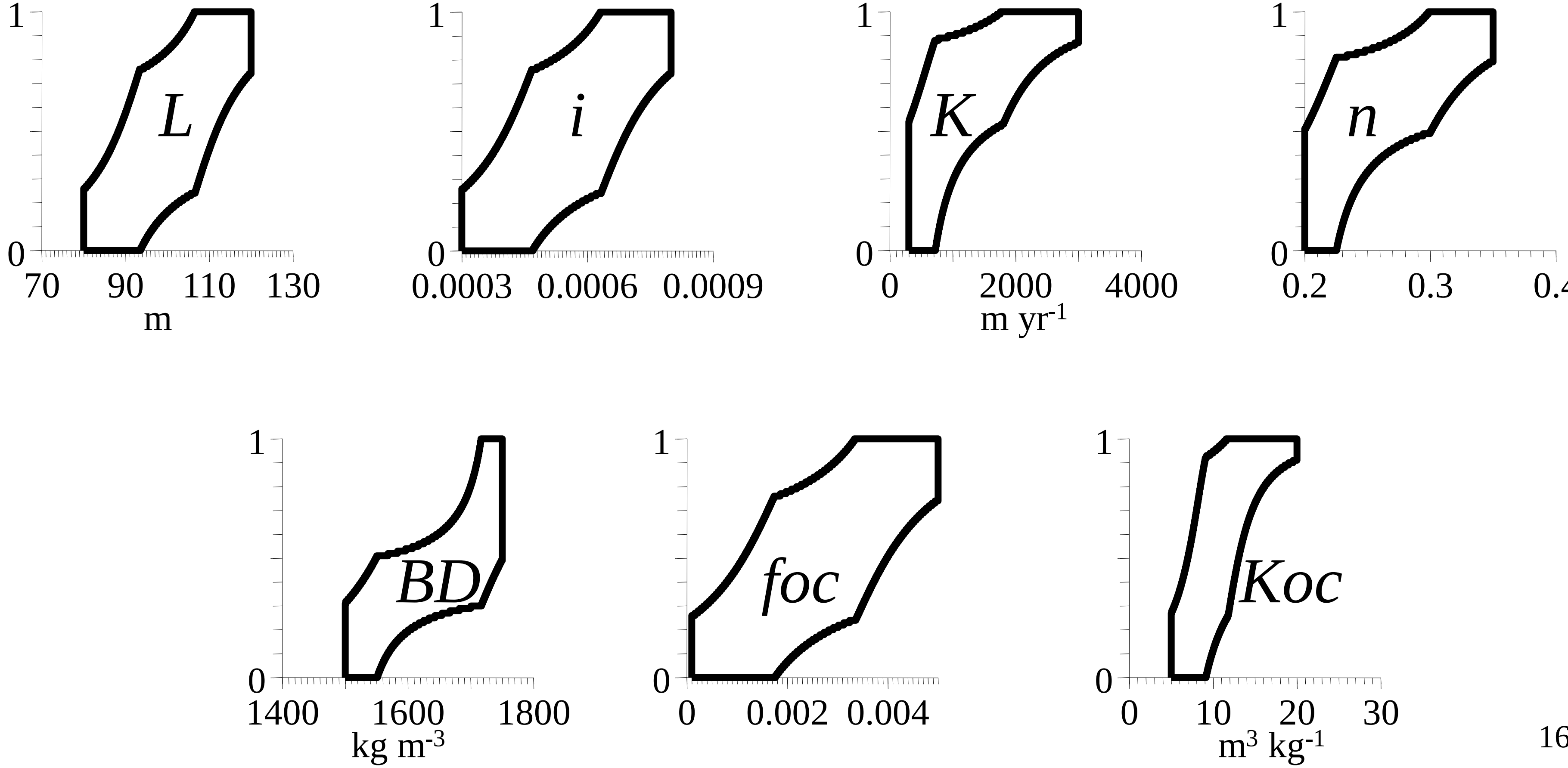


Example: Lobascio's travel time

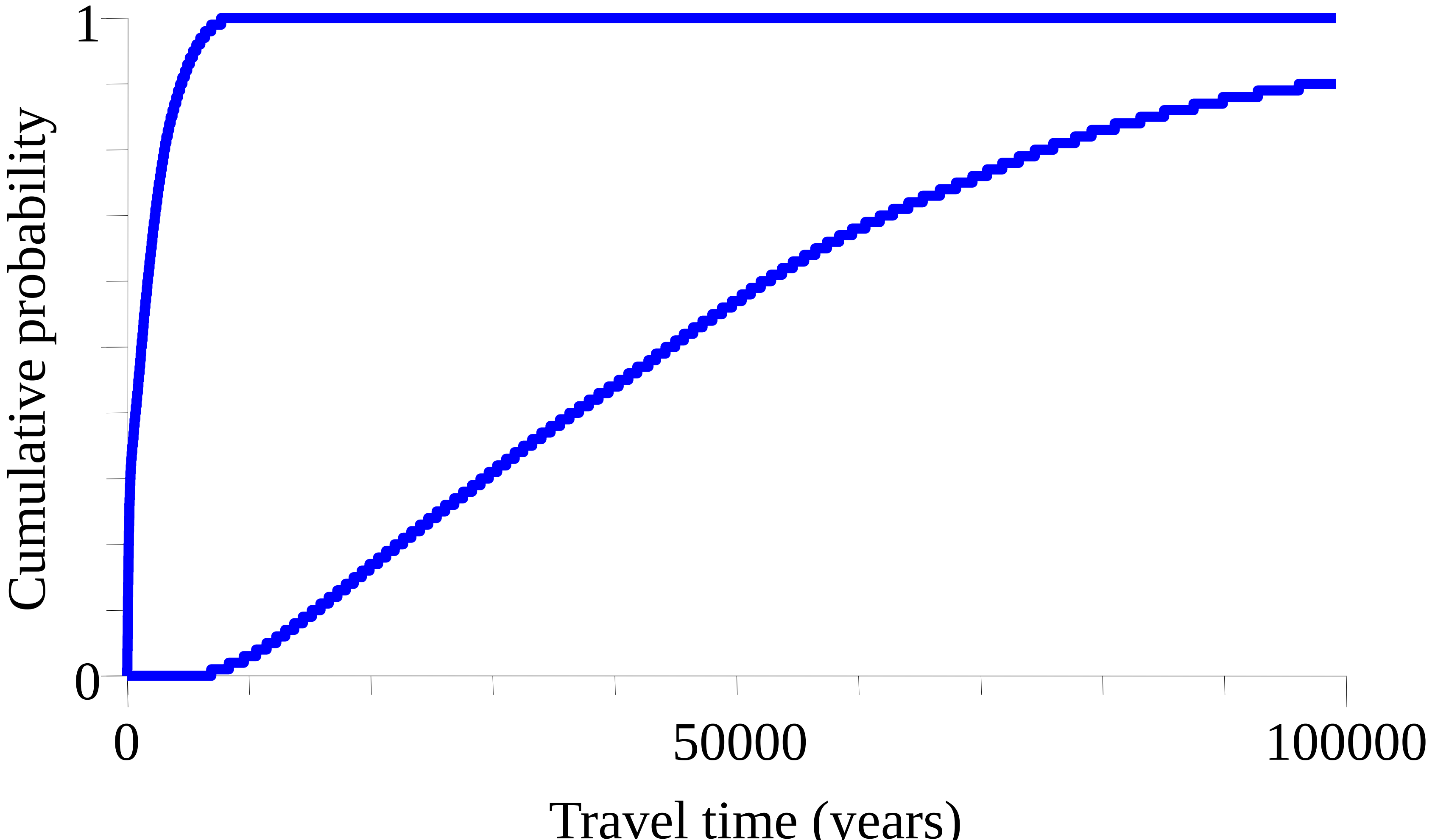
$$T = \frac{\left(n + BD \times foc \times Koc\right)L}{K \times i}$$

Parameter	Units	Min	Max	Mean	Stdv
L source-receptor distance	m	80	120	100	11.55
i hydraulic gradient	m/m	0.0003	0.0008	0.00055	0.0001443
K hydraulic conductivity	m/yr	300	3000	1000	750
n effective soil porosity	-	0.2	0.35	0.25	0.05
BD soil bulk density	kg/m^3	1500	1750	1650	100
foc fraction organic carbon	-	0.0001	0.005	0.00255	0.001415
Koc organic partition coefficient	m^3/kg	5	20	10	3

Inputs as mmms p-boxes

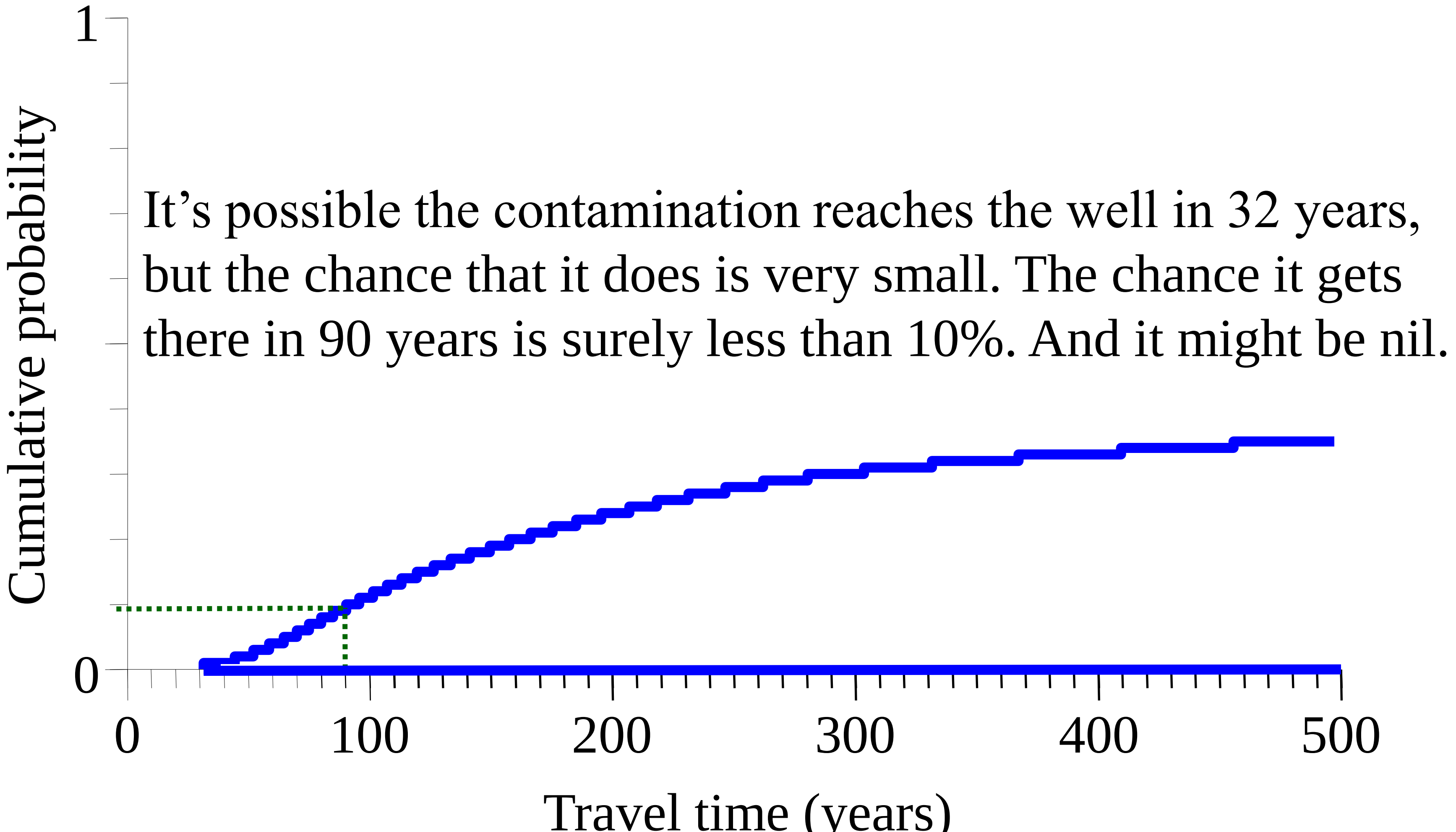


Output p-box



1
0
Cumulative probability
0
50000
100000
Travel time (years)

Detail of left tail



Example: mercury in wild mink

Location: Bayou d'Inde, Louisiana

Receptor: generic piscivorous small mammal

Contaminant: mercury

Exposure route: diet (fish and invertebrates)

Based loosely on the assessment described in "Appendix I2: Assessment of Risks to Piscivorus [sic] Mammals in the Calcasieu Estuary", *Calcasieu Estuary Remedial Investigation/Feasibility Study (RI/FS): Baseline Ecological Risk Assessment (BERA),* prepared October 2002 for the U.S. Environmental Protection Agency. See http://www.epa.gov/earth1r6/6sf/pdffiles/appendixi2.pdf.

Total daily intake from diet

$$TDI = FMR \times \left(\frac{C_{\mathrm{fish}} \times P_{\mathrm{fish}}}{AE_{\mathrm{fish}} \times GE_{\mathrm{fish}}} + \frac{C_{\mathrm{inverts}} \times P_{\mathrm{inverts}}}{AE_{\mathrm{inverts}} \times GE_{\mathrm{inverts}}} \right)$$

FMR	normalized free metabolic rate of the mammal
AE_{fish}	assimilation efficiency of dietary fish in the mammal
AE_{inverts}	assimilation efficiency of dietary invertebrates in the mammal
GE_{fish}	gross energy of fish tissue
GE_{inverts}	gross energy of invertebrate tissue
C_{fish}	mercury concentration in fish tissue
C_{inverts}	mercury concentration in invertebrate tissue
P_{fish}	proportion of fish in the mammal's diet
P_{inverts}	proportion of invertebrates in the mammal's diet

What is known about the variables

FMR, free metabolic rate
Studied in terms of mammal body mass; $FMR = a\, BW^{b}$, regression intervals a, b

BW, body mass of the mammal
Empirically well studied; Normal with highly precise mean and dispersion

AE_{fish}, AE_{inverts}, assimilation efficiencies for different diet components
A few field measurements; Mean and upper and lower values

GE_{fish}, GE_{inverts}, gross energy of fish and invertebrate tissues
Many measurements; Normal distribution with precise mean and dispersion

C_{fish}, C_{inverts}, mercury concentration in fish or invertebrate tissue
Dictated by EPA policy; Range between sample mean and 95% UCL

P_{fish}, P_{inverts}, proportions of fish and invertebrates in the mammal's diet
Assumed by analyst; Constant

Input assignments

BW = normal(608 gram, 66.9 gram)

FMR = [0.412±0.058] * *BW* ^ [0.862±0.026] * units('Kcal/kg/day')

*AE*fish = minmaxmean(0.77, 0.98, 0.91)

*AE*inverts = minmaxmean(0.72, 0.96, 0.87)

*GE*fish = normal(1200 Kcal per kg, 240 Kcal per kg)

*GE*inverts = normal(1050 Kcal per kg, 225 Kcal per kg)

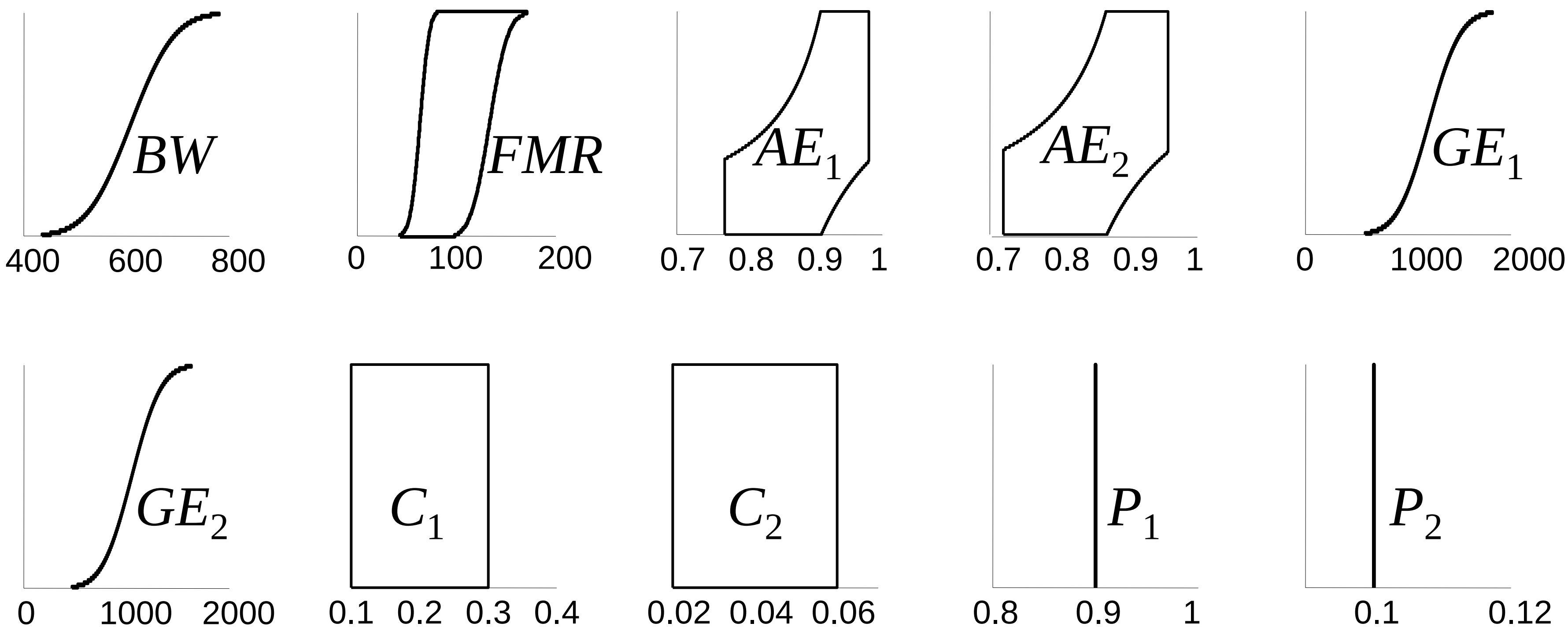
*C*fish = [0.1,0.3] mg per kg

*C*inverts = [0.02, 0.06] mg per kg

*P*fish = 0.9

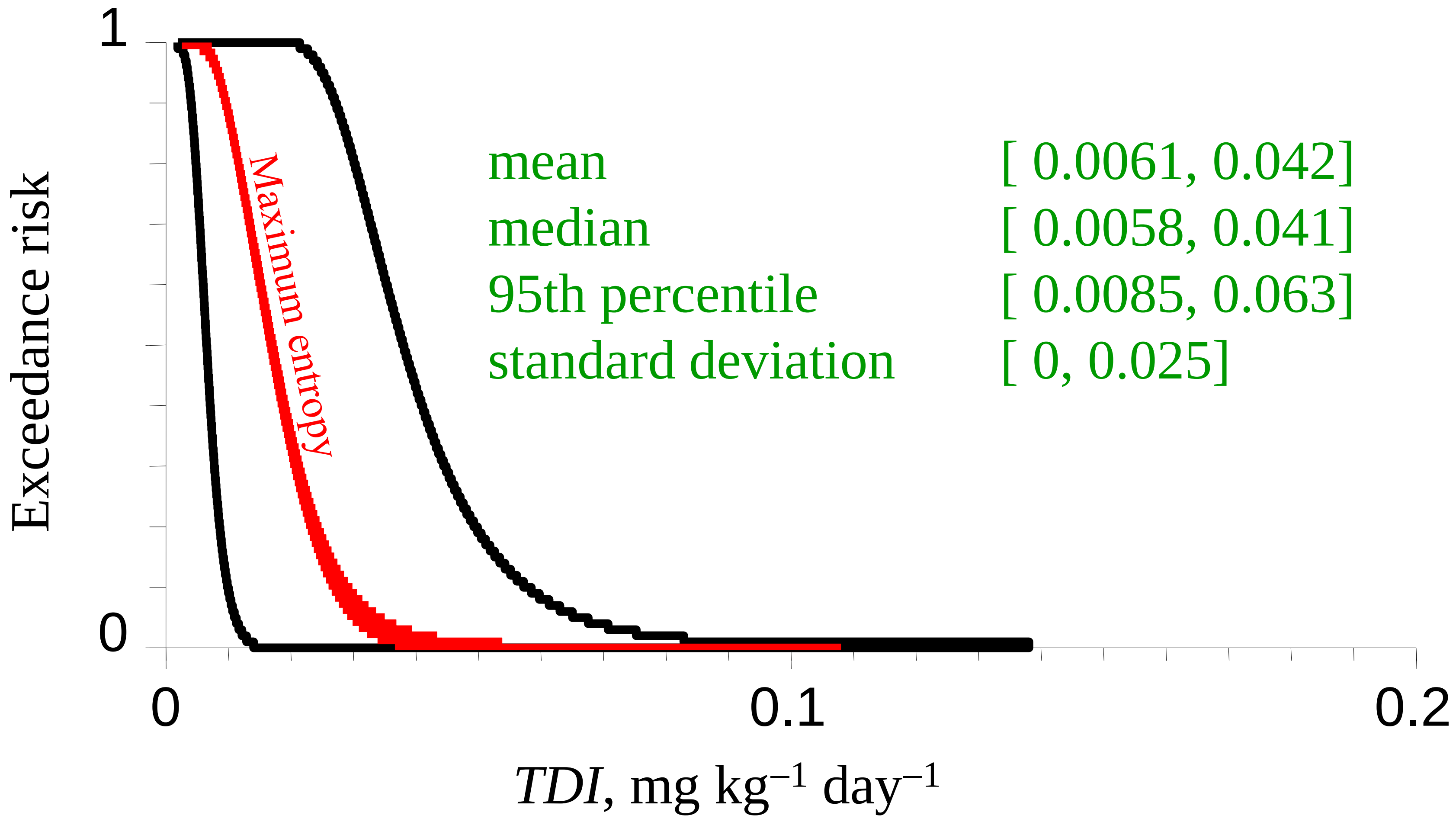
*P*inverts = 0.1

Input p-boxes



Subscript 1 denotes fish, 2 denotes inverts

Results



mean	[0.0061, 0.042]
median	[0.0058, 0.041]
95th percentile	[0.0085, 0.063]
standard deviation	[0, 0.025]

SKIP

Example: PCBs and duck hunters

Location: Berkshire County, Massachusetts

Receptor: Adult human hunters of waterfowl

Contaminant: PCBs (polychorinated biphenyls)

Exposure route: dietary consumption of contaminated waterfowl

Based on the assessment for non-cancer risks from PCB to adult hunters who consume contaminated waterfowl described in *Human Health Risk Assessment: GE/Housatonic River Site: Rest of River*, Volume IV, DCN: GE-031203-ABMP, April 2003, Weston Solutions (West Chester, Pennsylvania), Avatar Environmental (Exton, Pennsylvania), and Applied Biomathematics (Setauket, New York).

SKIP

Hazard quotient

$$HQ = \frac{EF \mid\times\mid IR \mid\times\mid C \mid\times\mid (1 \mid-\mid LOSS)}{AT \mid\times\mid BW \mid\times\mid RfD}$$

EF = mmms(1, 52, 5.4, 10) meals per year // exposure frequency, censored data, *n* = 23

IR = mmms(1.5, 675, 188, 113) grams per meal // poultry ingestion rate from EPA's *EFH*

C = [7.1, 9.73] mg per kg // exposure point (mean) concentration

LOSS = 0 // loss due to cooking

AT = 365.25 days per year // averaging time (not just units conversion)

BW = mixture(BW_{female}, BW_{male}) // body mass (Brainard and Burmaster 1992)

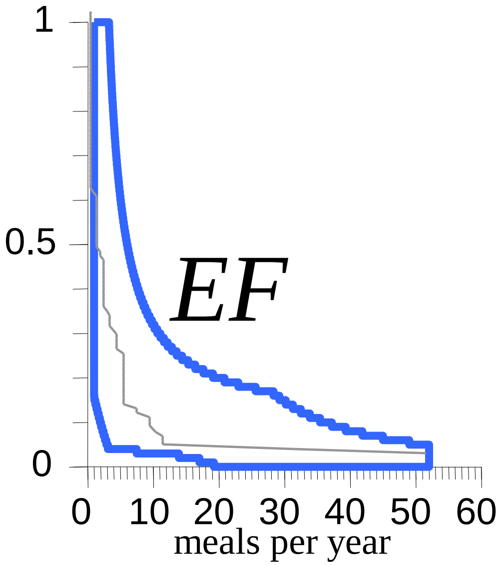
BW_{male} = lognormal(171, 30) pounds // adult male *n* = 9,983

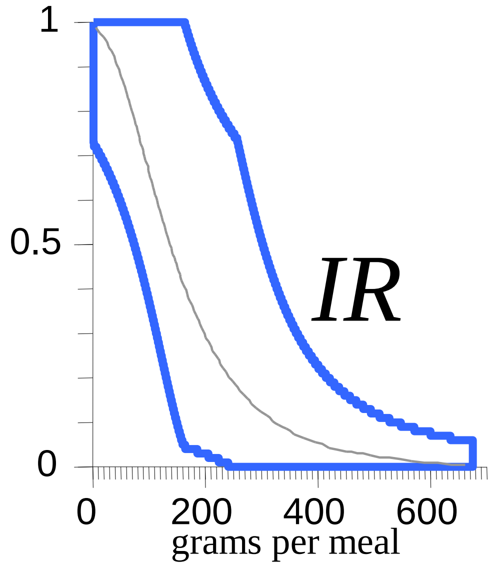
BW_{female} = lognormal(145, 30) pounds // adult female *n* = 10,339

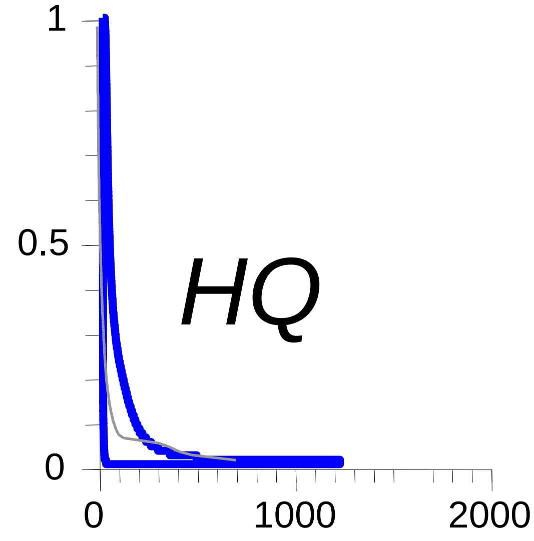
RfD = 0.00002 mg per kg per day // reference dose (EPA considers tolerable)

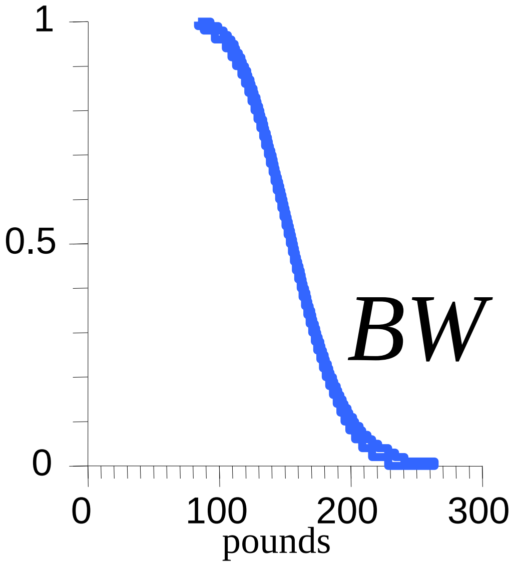
SKIP

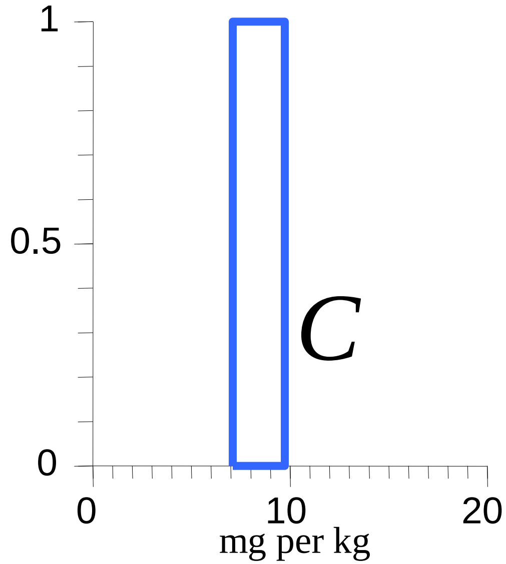
Inputs and results











mean	[14, 20]
standard deviation	[32, 47]
median	[1, 36]
95th percentile	[5, 225]
range	[0.01, 1230]

SKIP

2MC simulations may not fill p-boxes

- 2^{nd} order Monte Carlo is not comprehensive
  - Inadequate model of ignorance
  - Dependence among parameters of a distribution
  - Uncertainty about dependence (Fréchet)
  - Non-denumerable model uncertainty
- Probability bounds analysis is not optimal
  - Independence between parameters of a distribution
  - Ternary (and higher) Fréchet operations

SKIP

Probability bounds analysis

- Combines interval and probability methods
- Shows when uncertainty is (or isn't) important
- Won't underestimate risks of extremes, yet isn't hyperconservative like worst case
- Solves many risk analysis problems
  - Input distributions unknown
  - Large measurement error, censored data, and small samples
  - Correlation and dependency ignored
  - Model uncertainty

Advantages

- Fewer assumptions
  - Not just *different* assumptions, *fewer* of them
  - Distribution-free methods
- Rigorous results
  - Built-in quality assurance
  - Automatically verified calculation

What p-boxes can't do

- Show what's most likely within a p-box
- Express second-order information
- Get best-possible bounds on non-tail risks
- Get best-possible bounds for any data set
  - When dependencies are intricate
  - When information about modes are availabe

Don't know the input distributions

Solved

- Don't *have* to specify the distributions
- Shouldn't use a distribution without evidence
- Maximum entropy criterion erases uncertainty rather than propagates it
- Sensitivity analysis is very hard since it's an infinite-dimensional problem
- P-boxes easy, but should use all information

Install the pba.r probability box library

- Invoke R by double clicking its desktop icon
- Enter **rm(list = ls())** to clear R's memory
- Click File / Source R code on the main menu
- Locate and Open the file pba.r
- You'll get the message "`:pba> library loaded`"

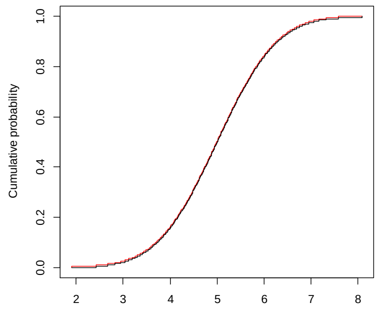
Die-hard RStudio users

- If, against advice, you must use RStudio…
- File / Open file `pbox.r`
- Source (Ctrl+Shift+S) `pbox.r`
- Enter the instruction **RStudio = TRUE**

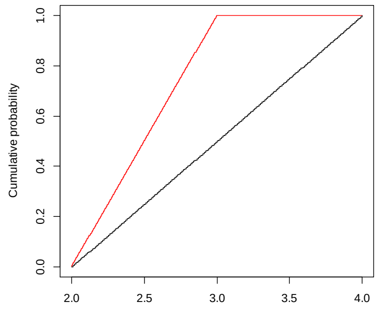
pba.r probability bounds library

a = normal(5,1)

a

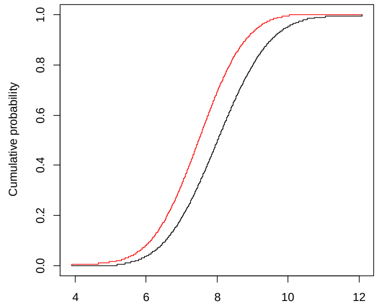


b = uniform(2, interval(3,4))



b

a + b



Generalized convolutions

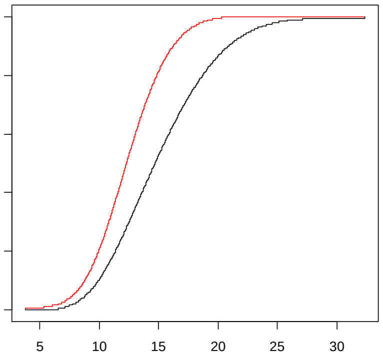
a * b

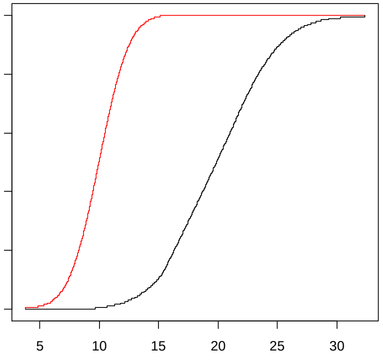
a %*% b

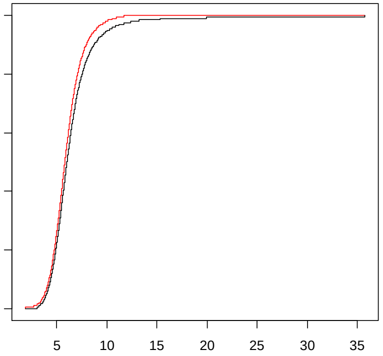
c = a + weibull(8,1) * beta(2,3) / b

c

mean(c)







Probability bounds analysis in pba.r

- Output
  - <enter variable name>, plot, lines, show, summary
- Characterize
  - mean, sd, var, median, quantile, fivenum, left, right, prob, cut, percentile, iqr, random, range
- Compute
  - exp, log, sqrt, abs, round, trunc, ceiling, floor, sign, sin, cos, tan, asin, acos, atan, atan2, reciprocate, negate, +, -, *, /, pmin, pmax, ^, and, or, not, mixture, smin, smax, complement

Probability bounds analysis in pba.r

- Construct
  - <named distribution>
  - histogram
  - quantiles
  - pointlist
  - MM <tab> <tab>
  - ME <tab> <tab>
  - ML <tab> <tab>
  - CB <tab> <tab>
  - NV <tab> <tab>

Supported named distributions

bernoulli, beta (B), binomial (Bin), cauchy, chi, chisquared, delta, dirac, discreteuniform, exponential, exponentialpower, extremevalue, F, f, fishersnedecor, fishertippett, fisk, frechet, gamma, gaussian, geometric, generalizedextremevalue (GEV), generalizedpareto, gumbel, histogram, inversegamma, laplace, logistic, loglogistic, lognormal (L), logtriangular, loguniform, negativebinomial, normal (N), pareto, pascal, powerfunction, poisson, quantiles, rayleigh, reciprocal, shiftedloglogistic, skewnormal (SN), student, trapezoidal, triangular (T), uniform (U), weibull

Neuroscience of risk

Decade of the Brain ('90s)

Psychometry

- Probability and decision theory are rife with paradoxes that no other areas in math have
- Cottage industry in documenting ways in which humans mess up probabilities

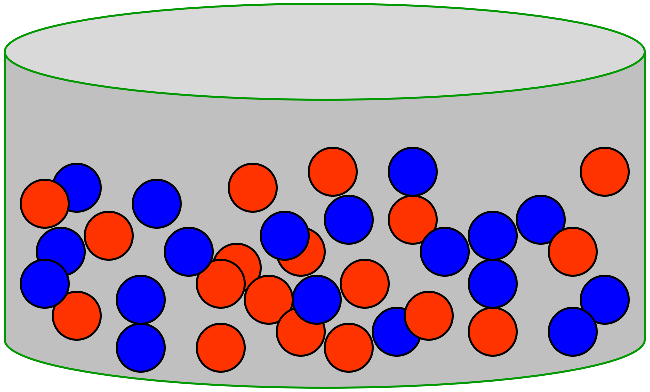
Paradoxes & biases

- Ellsberg paradox
- St. Petersburg paradox
- Two-envelopes problem
- Monty Hall problem
- Simpson’s paradox
- Risk aversion
- Loss aversion

Risk aversion

- Suppose you can get $100 if a randomly drawn ball is red from an urn with half red and half blue balls…or you can just get $50
- Which prize do you want?

$50

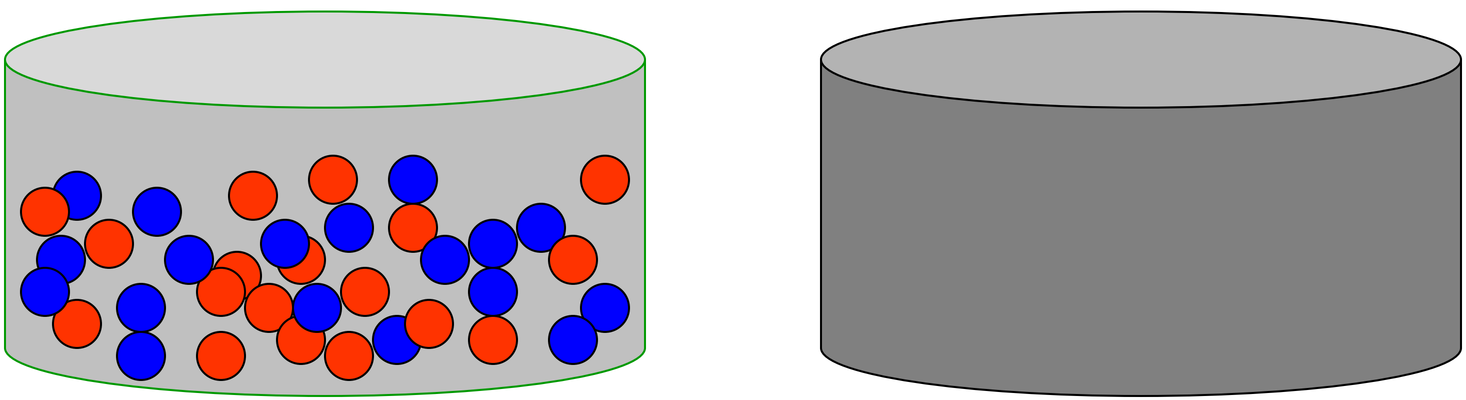


EU is the same, but most people take the sure $50

Ambiguity aversion

Keynes; Dempster

- Balls can be either red or blue
- Two urns, both with 36 balls
- Get $100 if a randomly drawn ball is red
- Which urn do you wanna draw from?



Ellsberg Paradox

- Balls can be red, blue or yellow (probs are R, B, Y)
- A well-mixed urn has 30 red balls and 60 ***other*** balls
- Don't know how many are blue or how many are yellow

Gamble A $R > B$	**Gamble B**
Get \$100 if draw red	Get \$100 if draw blue
Gamble C	**Gamble D** $R < B$
Get \$100 if red or yellow	Get \$100 if blue or yellow

Persistent paradox

- People always prefer unambiguous outcomes
  - Doesn’t depend on your utility function or payoff
  - Not related to risk aversion
  - We simply don’t like ambiguity
- Not explained by probability theory, or by prospect theory

Ambiguity (incertitude)

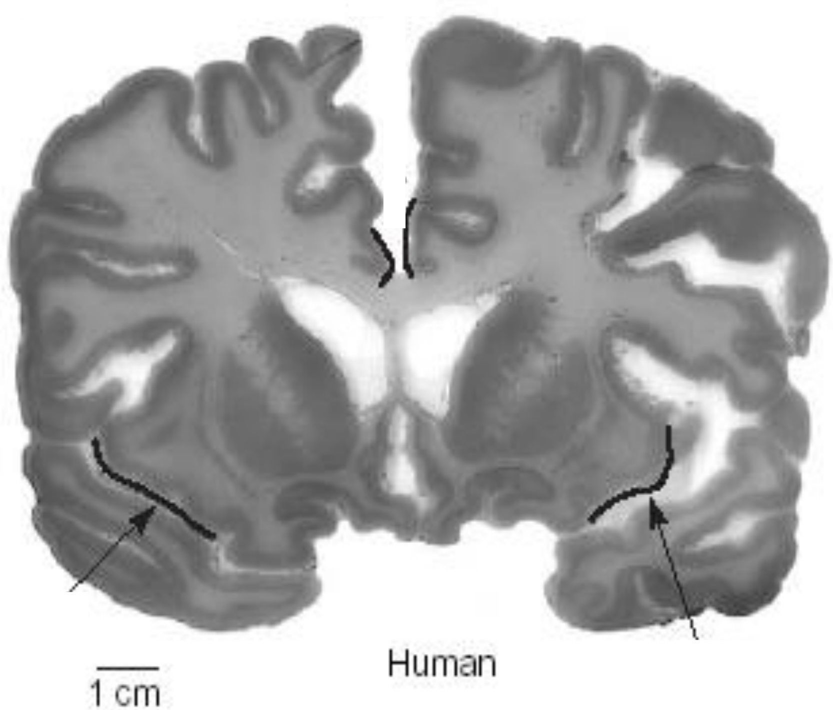
- Ambiguity aversion is ubiquitous in human decision making, and is utterly incompatible with Bayesian norms

- Humans are *wired* to process incertitude separately and differently from variability

Neuroscience of risk perception

(Marr 1982; Barkow et al. 1992; Pinker 1997, 2002)

Instead of being divided into rational and emotional sides, the human brain has *many* special-purpose calculators



Partial list of mental calculators

(after Pinker 2002; 1997; Marr 1982; Barkow et al. 1992)

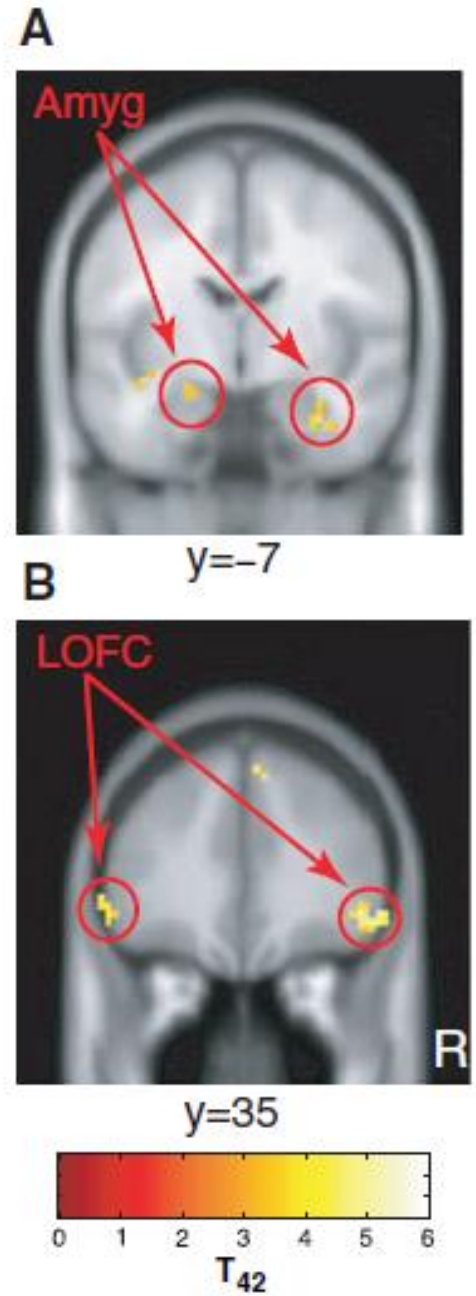
- Language (grammar and memorized dictionary)
- Practical physics (pre-Newtonian)
- Intuitive biology (animate differs from inanimate)
- Intuitive engineering (tools designed for a purpose)
- Spatial sense (dead reckoner and mental maps)
- Number sense (1, 2, 3, many)
- Probability sense (frequentist Bayes)
- Uncertainty detector (procrastination)
- Intuitive economics (reciprocity, trust, equity, fairness)
- Intuitive psychology (theory of mind, deception)

What's the evidence for an 'uncertainty detector' in humans?

fMRI

- Hsu et al. (2005) found localized regions of activity in the brain under situations of ambiguity (incertitude)

- Amygdala associated with processing fear and threat



Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310: 1680-1683.

Ambiguity/incertitude detector

- Humans have an incertitude processor
  - Triggered by situations with ambiguity
  - Especially focused on the worst case
  - Common response is procrastination
- Functional organ
  - Normal feature of the human brain, also apes, rats
  - Not a product of learning
  - Visible in fMRI
- Brain lesions can make people insensitive to incertitude…so they behave as Bayesians

Other species

- Chimpanzees and bonobos preferred peanuts (which they like less than bananas) when they don’t know the probability of getting bananas

Rosati, A., and B. Hare 2010. Chimpanzees and bonobos distinguish between risk and ambiguity. *Proceedings of Royal Society: Biology Letters*. See also “Apes unwilling to game when odds are uncertain” http://www.physorg.com/print209830622.html

What's the evidence for a 'probability sense' in humans?

Probability sense

- One of the ways we know there is a probability calculator is that we can watch it turn on
- Platt and Glimcher (NYU) found particular neurons in the lateral intraparietal cortex in rhesus monkeys encode both the probability of an outcome and its magnitude
- We can also see it in reasoning behaviors

Bayesian reasoning (poor)

If a test to detect a disease whose prevalence is 0.1% has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming that you know nothing about the person's symptoms or signs? ___%

12-18% of medical students get this correct

Casscells et al. 1978 replicated in Cosmides and Tooby 1996

Bayesian reasoning (good)

If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 50/1000, what is the chance that a person found to have a positive result actually has the disease, assuming that you know nothing about the person's symptoms or signs? __**1**__ out of __**51**__.

8 or 9 in 10 medical students get this correct.
That it's so easy to solve suggests hardwiring.

Casscells et al. 1978 replicated in Cosmides and Tooby 1996

Calculators must be triggered

Cosmides & Tooby 1996; Gigerenzer 1991

- Humans have an innate probability sense, but it is triggered by *natural frequencies*

  Format of sensory data matters, not the meaning

- This calculator kicked in for the medical students who got the question in terms of natural frequencies, and they mostly solved it

- The mere presence of the percent signs in the question hobbled the other group

Multiple calculators may fire

- There are distinct calculators associated with
  - Probabilities and risk (variability) medical students
  - Ambiguity and uncertainty (incertitude) Hsu et al.
  - Trust and fairness Ultimatum Game
- Brain processes them differently
  - Different brain regions
  - Different chemical systems
- They can give conflicting answers
- (e.g. Glimcher & Rustichini 2004 and references therein)

Conflict explains Ellsberg paradox

- Ambiguity detector countermands any risk estimate that might be produced by the probability sense
- What matters most, and perhaps exclusively, is *how bad it could be*…not how likely or unlikely that outcome is

Biological basis for Ellsberg

Hsu et al. 2005

- Probability sense and the ambiguity detector interfere with each other
- Humans do not make decisions based purely on probability in such cases
- Probabilists use equiprobability to model incertitude which confounds it with variability

Conflict may explain other biases too

Kahneman and Tversky

- Ambiguity aversion
  Avoiding options when probabilities seem unknown
- Loss aversion
  Disliking a loss more fervently than liking a gain of the same magnitude
- Hyperbolic discounting
  Preferring immediate payoffs over later (more intense the closer to present payoffs are)
- Base rate fallacy
  Neglecting available statistical data in favor of particulars
- Neglect of probability
  Disregarding probability in decision making under uncertainty
- Pseudocertainty
  Making risk-averse choices for positive outcomes, but risk-seeking for negative

People just seem downright *stupid* about risks and uncertainty

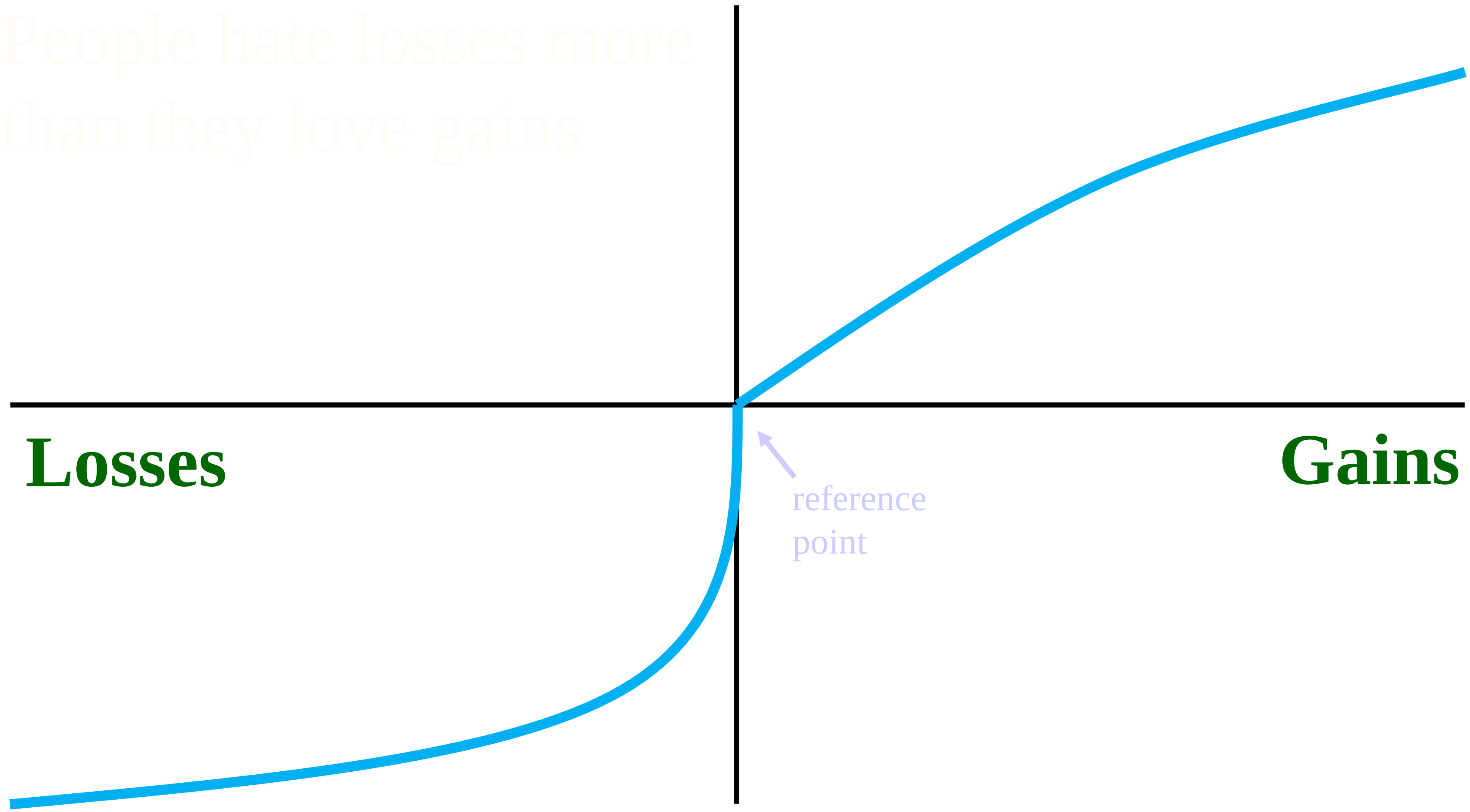
SKIP

Loss aversion

(asymmetry in perceptions about losses and gains)

SKIP

Loss aversion



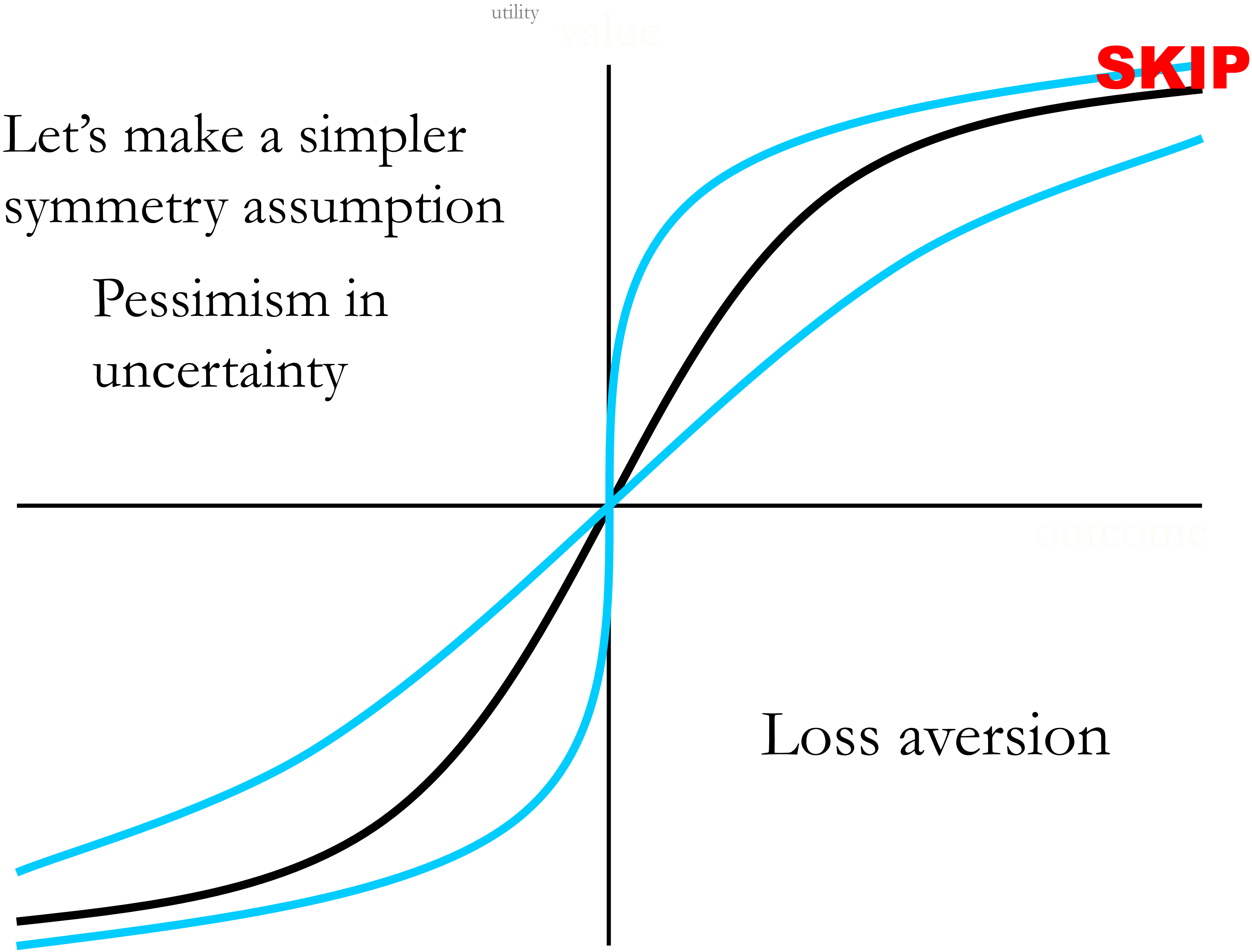
Prospect theory adopts (but does not explain) loss aversion

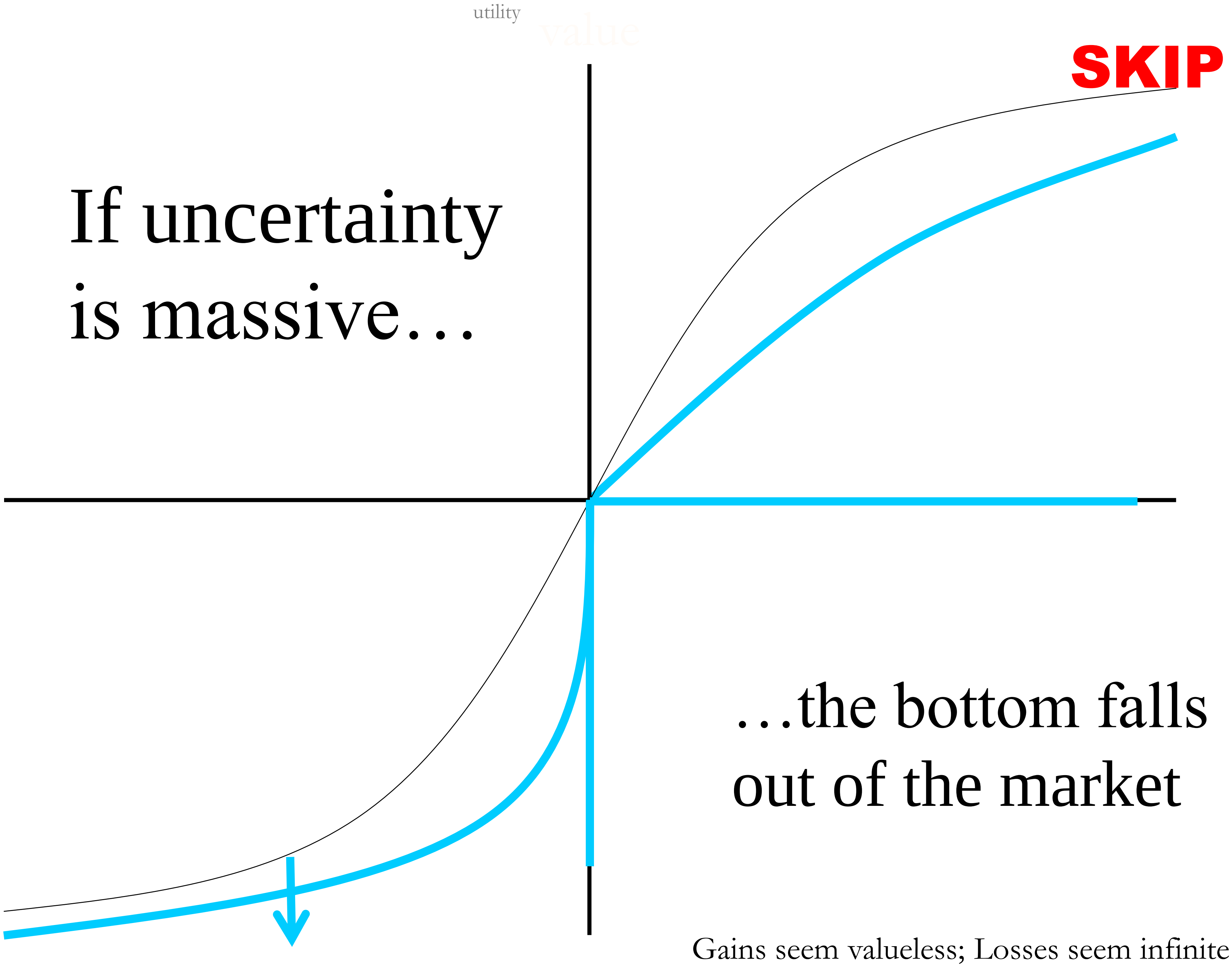
SKIP

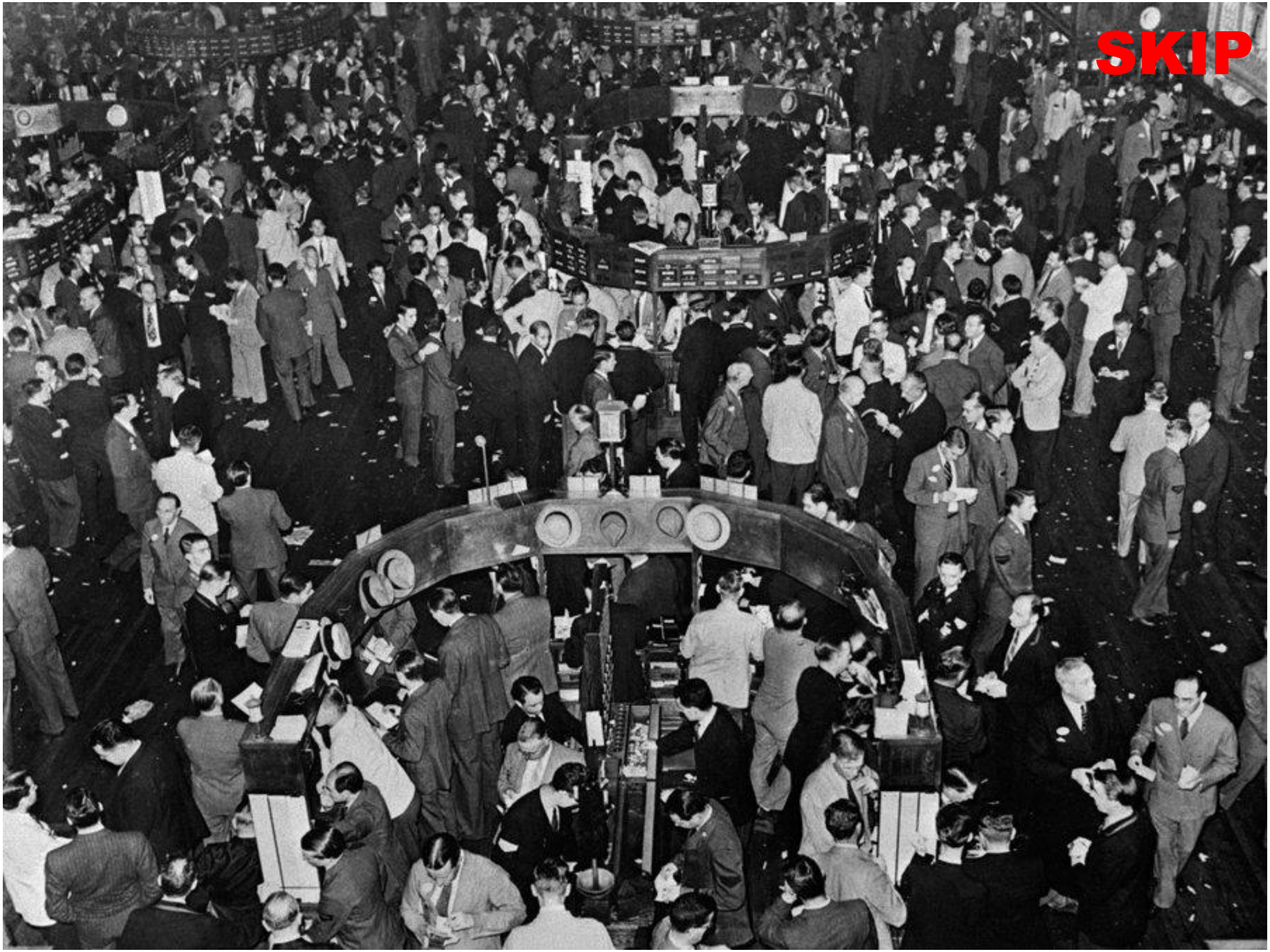
But *why*?

- Prospect theory is the state of the art
- Purely descriptive
- Doesn't say *why* loss aversion should exist

- What is the biological basis for loss aversion?
- How could it have arisen in human evolution?


utility
SKIP
Let's make a simpler symmetry assumption
Pessimism in uncertainty
Loss aversion


utility
SKIP
If uncertainty is massive…
…the bottom falls out of the market
Gains seem valueless; Losses seem infinite


SKIP

SKIP

Loss aversion disappears with certainty

- Loss aversion disappears
  - with a person you trust, or
  - *after* the gamble has been realized
    - Gilbert et al. 2004
    - Kermer et al. 2006
    - Yechiam & Ert 2007
    - Erev, Ert, & Yechiam 2008
    - Ert & Erev 2008
- When losses and gains are surely exchangeable, the uncertainty contracts to the symmetric utility

SKIP

Direct experimental evidence

- Ellsberg made the probabilities ambiguous
- Psychologist Christian Luhmann (Stony Brook) made *rewards* ambiguous
  - Visually obscured the promised payoffs
  - "I'll pay you between 1 and 10 bucks"
- Loss aversion varies with the size of uncertainty
- Disappears with certainty

SKIP

Clinical evidence

- Amygdala damage eliminates loss aversion
- But doesn't affect a person's ability to gamble and respond to changing values and risk (n = 2)

Still normal in risk aversion

- Amygdalectomied rhesus monkeys approach stimuli that healthy monkeys avoid

De Martino, B., C.F. Cramerer and R. Adolphs (2010). Amygdala damage eliminates monetary loss aversion. *Proceedings of the National Academy of Sciences of the United States of America* 107(8): 3788–3792.
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2840433/pdf/pnas.200910230.pdf
Mason et al. (2006). *Emotion* 6: 73-81.

SKIP

But why pessimism?

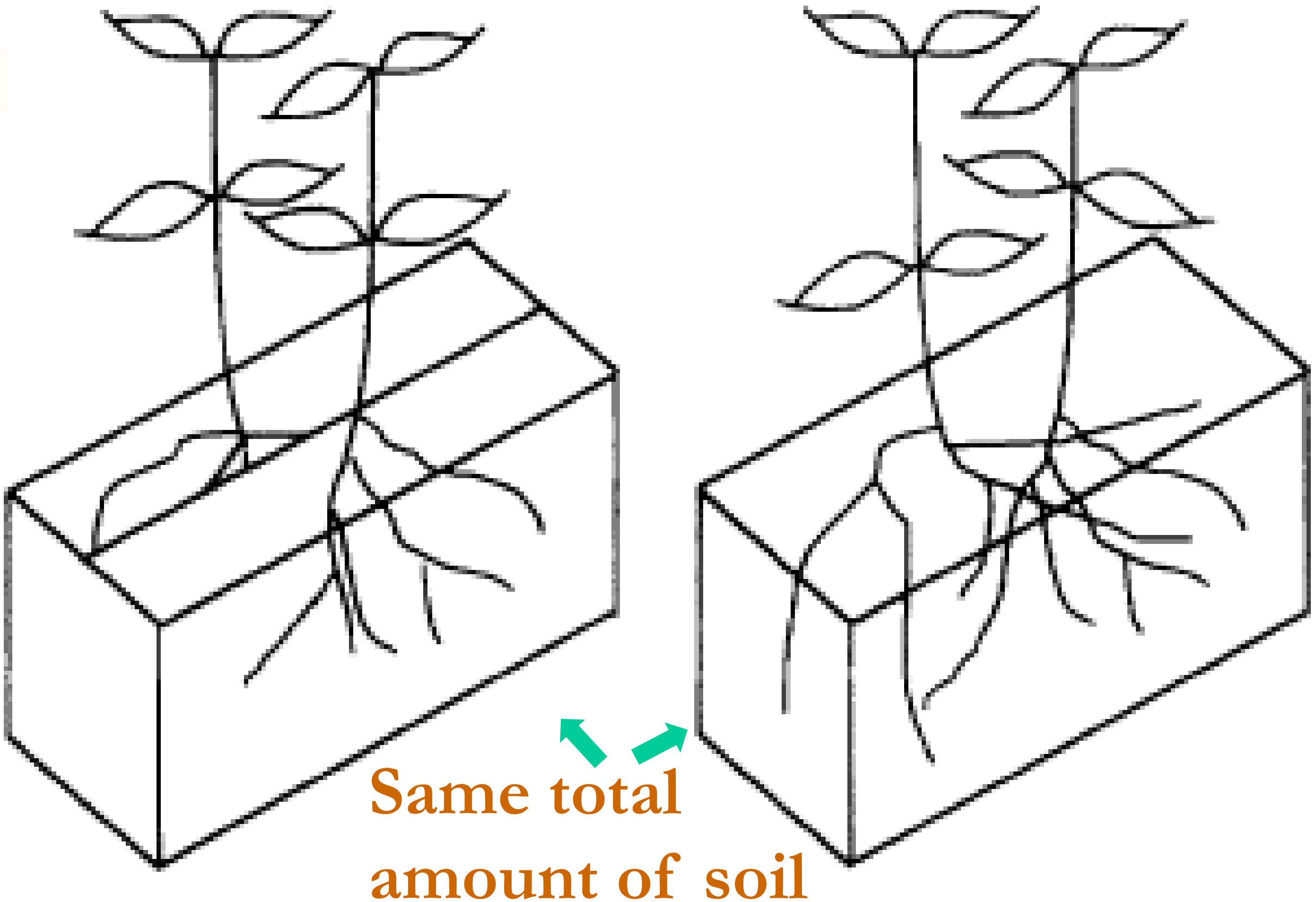
- Pessimism is often advantageous evolutionarily

- Natural selection can favor pessimism
  - Death is ‘hard selection’
  - Animal foraging strategies
  - Programmed plant behaviors

- Being wrong often has asymmetric consequences
  - Foraging: Finding dinner versus being dinner
  - Competition: Preemption versus being preempted

And even in plants!

SKIP



Gersani, M., Brown, J. S., Brien, E. E., Maina, G. M., & Abramsky, Z. (2001). Tragedy of the commons as a result of root competition. *Journal of Ecology* 89: 660-669.

SKIP

Pessimism is not inevitable

- Pessimism is not the only reaction to uncertainty
  - Normal people in stressful situations
  - Pathological gamblers
  - Maniacs

- Ambiguity aversion decreases with optimism (Pulford 2009)

Pulford, B. D. (2009). Is luck on my side? Optimism, pessimism, and ambiguity aversion. *Quarterly Journal of Experimental Psychology*, 62: 1079-1087.

Third calculator: intuitive economics

- Computes fairness of situations
- Detects cheaters who are getting more than their share, or not shouldering their responsibility

Ultimatum Game: a scientist offers money to two players

- One player proposes how the money should be divided between the two players
- The other player can *accept* (and both get their share) or *reject* the division (and neither gets anything)

How should you play?

- Economists say the rational behavior is
  - Responder: always accept any deal (it's free money!)
  - Proposer: always offer the smallest possible amount
- Actual plays are much closer to fair
  - Proposer offers a split much closer to 50:50
  - Responder accepts only if the split is closer to 50:50

Universal in humans

- Ultimatum Game is only surprising to economists
- The pattern may be universal in human behavior
  - Over 100 papers in 25 western societies
  - 15 non-western societies
- Exceptions
  - Sociopaths
  - Chimpanzees
  - Kids under 5
  - People playing against machines
- Chimps are *more rational* that humans

Why do humans do this?

- Adaptation for reciprocal altruism
  - Cares about fairness and reciprocity
  - Alters outcomes of others at a personal cost
  - Rewards those who act in a prosocial manner
  - Punishes those who act selfishly, even when punishment is costly
- Mediated by the fairness calculator
  - Pattern absent when fairness is not an issue

“Irrationality”

- Irrationality is a hallmark of human decisions

- Why are humans biased, irrational, stupid?
  - Using the wrong mental calculator *(optical illusion)*
  - Disagreement among mental calculators
  - Concerned with issues outside the risk analysis
    - Fairness, justice
    - Outcomes not treated in the analysis
    - Chance the risk analyst is lying
    - Chance the risk analyst is inept

Different calculators

Currently, confusion is guaranteed

- Neuroimagery and clinical psychology show humans distinguish incertitude and variability

- Probabilists traditionally use equiprobability to model incertitude, which confounds the two

- Risk analysts report their findings in ways that we know will create misunderstandings

Import for risk assessment

- Risk analyses woefully incomplete
  - Neglect or misunderstand incertitude
  - Omit important issues and thus understate risks
- Presentations use very misleading formatting
  - Percentages, relative frequencies, conditionals, etc.
- Both problems can be fixed
  - By changing *analysts'* behavior (not the public's)

Risk communication

- Informed consent in medicine
- Effective communication to decision makers

- If you want people to understand your risk calculations, you have to *speak their language*

Take-home messages

- Evolution has wired humans to see incertitude distinctly, and differently, from variability
- Conflict between the two seems to explain many probability and decision paradoxes
- There is a proper calculus that handles both although it is incompatible with current norms

Wishful thinking

In practice, risk analysts often use conventions and make assumptions that may be convenient but are not really justified:

1. All variables are independent of one another
2. Uniform distributions model incertitude
3. Distributions are stationary (unchanging)
4. Specifications are perfectly precise

The assumptions may understate risks

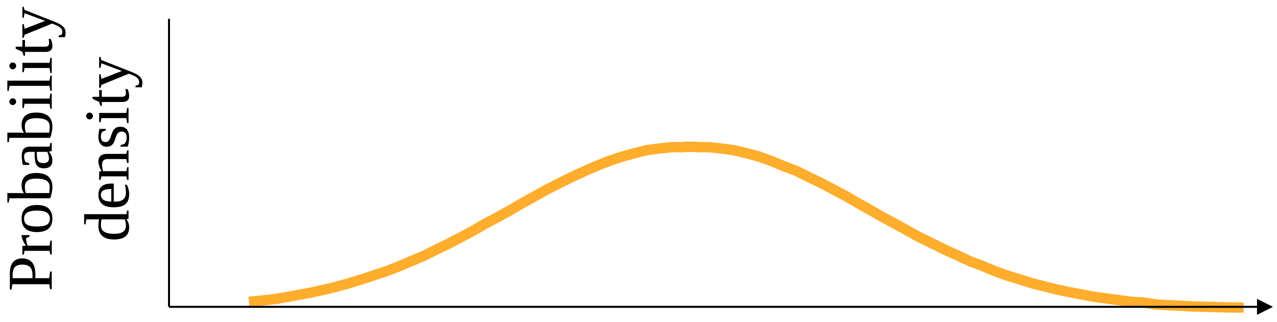
- Probability theory, as its commonly used, doesn't cumulate gross uncertainty correctly

- Precision of the answer (measured as cv) depends strongly on the number inputs and not so strongly on their distribution shapes, even if they are uniforms or flat priors

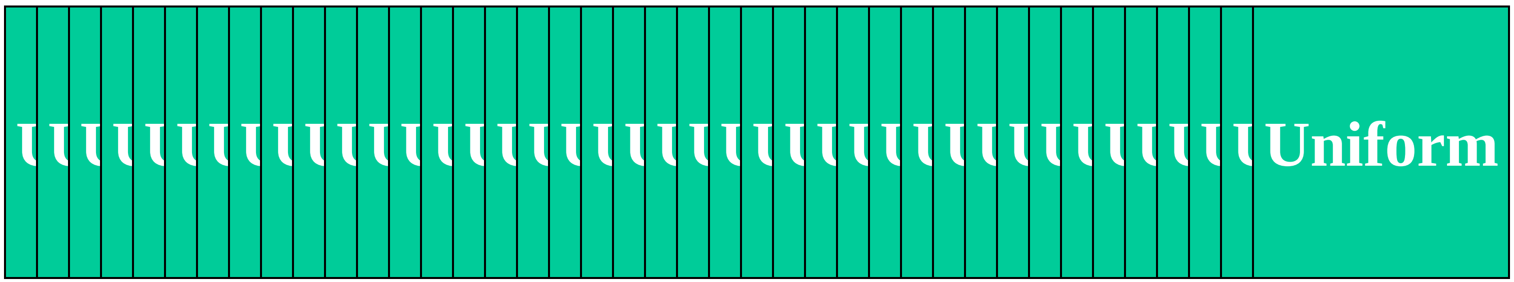
- The more inputs, the tighter the answer

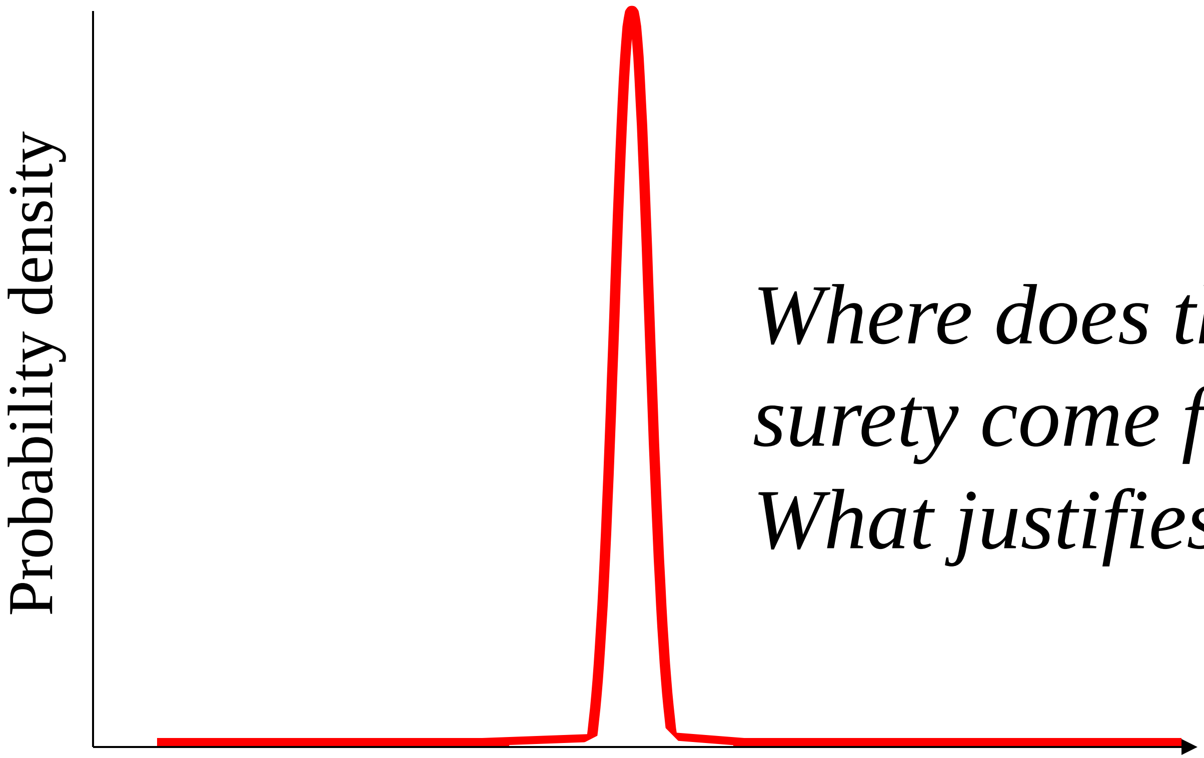
A few grossly uncertain inputs

Uniform + **Uniform** + **Uniform** + **Uniform**



A lot of grossly uncertain inputs...





Where does this surety come from? What justifies it?

Smoke and mirrors certainty

- P-boxes give a vacuous answer if all you give them are vacuous inputs

- Conventional probability theory, at least as it's naively applied, seems to manufacture certainty out of nothing

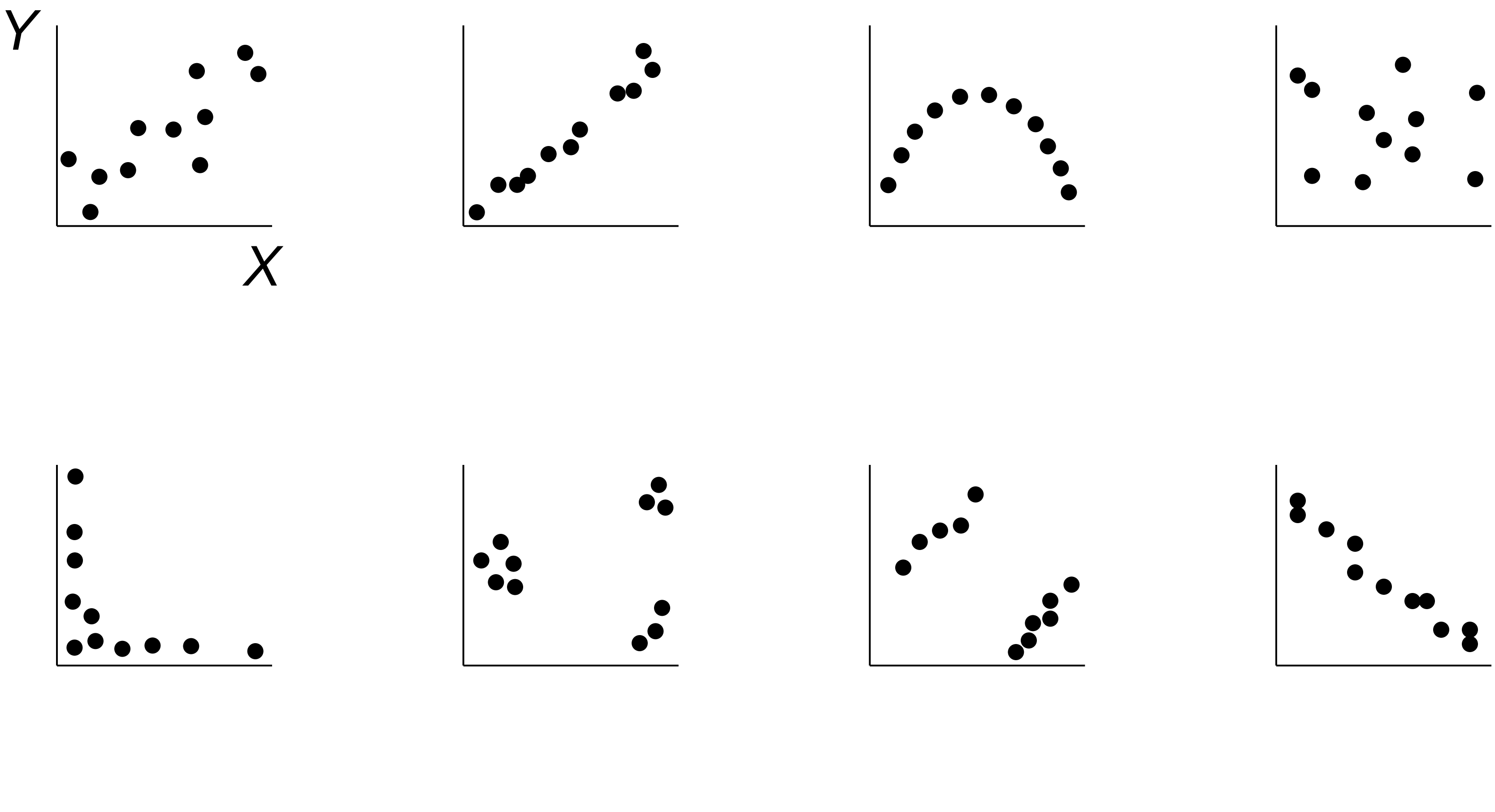
- This is why some critics say probabilistic analyses are "smoke and mirrors"

Correlation and dependency

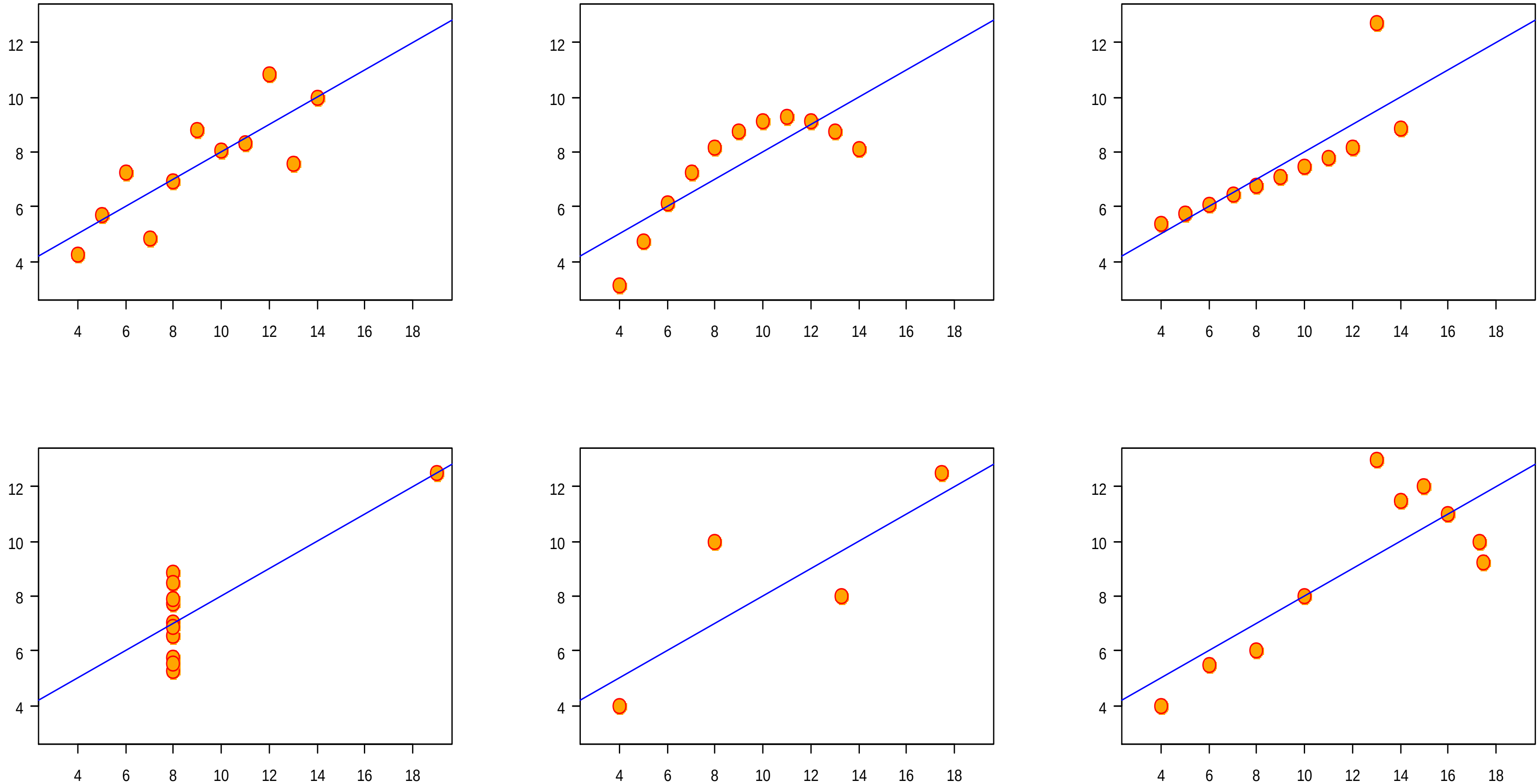
Dependence

- Most variables assumed independent
- But some variables clearly aren't
  - Density and porosity
  - Economic status and general health
  - Body size and skin surface area

Dependence can be complex



Even for fixed correlation



Every scattergram has (Pearson) correlation 0.816

SKIP

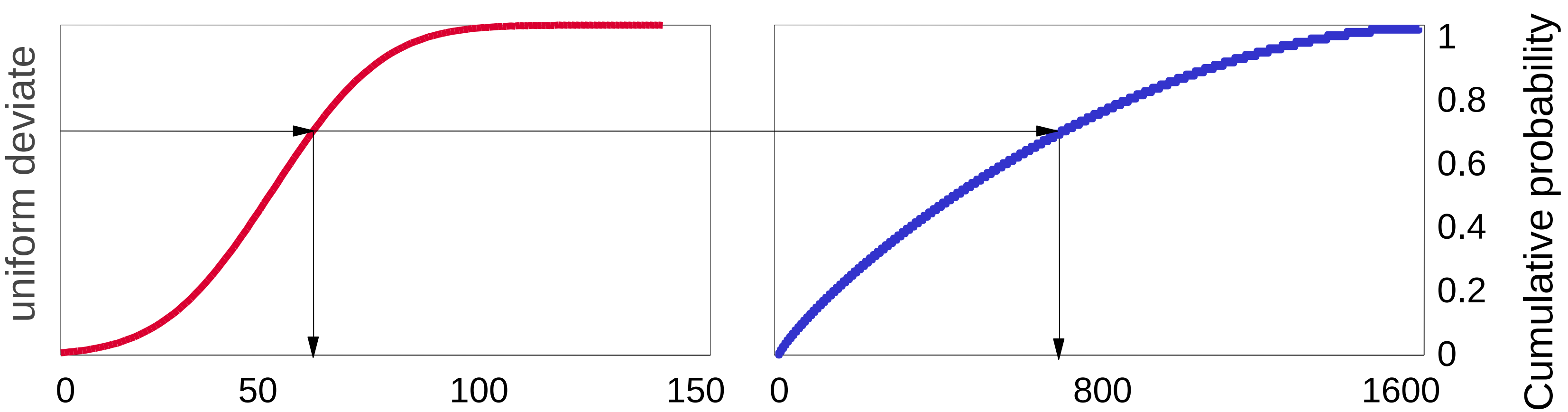
Dependencies

- Independence (knowing X tells nothing about Y)
- Perfect dependence $F(X) = G(Y)$
- Opposite dependence $F(X) = 1-G(Y)$
- Complete dependence $Y = z(X)$
- Linearly correlated $Y = mX + b + \varepsilon$
- Ranks linearly correlated $G(Y) = mF(X) + b + \varepsilon$
- Functional modeling $Y = z(X) + \varepsilon$
- Complex dependence (anything else!)

Uncorrelatedness is not independence

Perfect dependence

- May be better than assuming independence
- Very easy to simulate



SKIP

Specified correlations

$$Z_i = \sum_{j=1}^{i} c_{ij} Y_j , \quad i = 1,\ldots,k$$

- Scheuer and Stoller (1962)
- Pearson, not rank, correlation
- Only for *normal* marginals

$$\Sigma = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1k} \\ \vdots & \ddots & \vdots \\ \sigma_{k1} & \cdots & \sigma_{kk} \end{bmatrix} = CC^T$$

Correlated deviates are weighted sums of independent standard normal deviates Y_j ~ normal(0,1) where weights c_{ij} are the elements of a lower triangular matrix C solving $\Sigma = CC^{\mathrm{T}}$ (by Cholesky decomposition) where Σ is the intended variance-covariance matrix (in which diagonals are variances off-diagonals are covariances)

SKIP

Bivariate case

If Y_1, Y_2 ~ normal(0,1) are independent, then

$$Z_1 = Y_1\sigma_1 + \mu_1$$
$$Z_2 = \left(rY_1 + \sqrt{1 - r^2}\,Y_2\right)\sigma_2 + \mu_2$$

will produce deviates such that

$Z_1 \sim \text{normal}(\mu_1, \sigma_1)$

$Z_2 \sim \text{normal}(\mu_2, \sigma_2)$

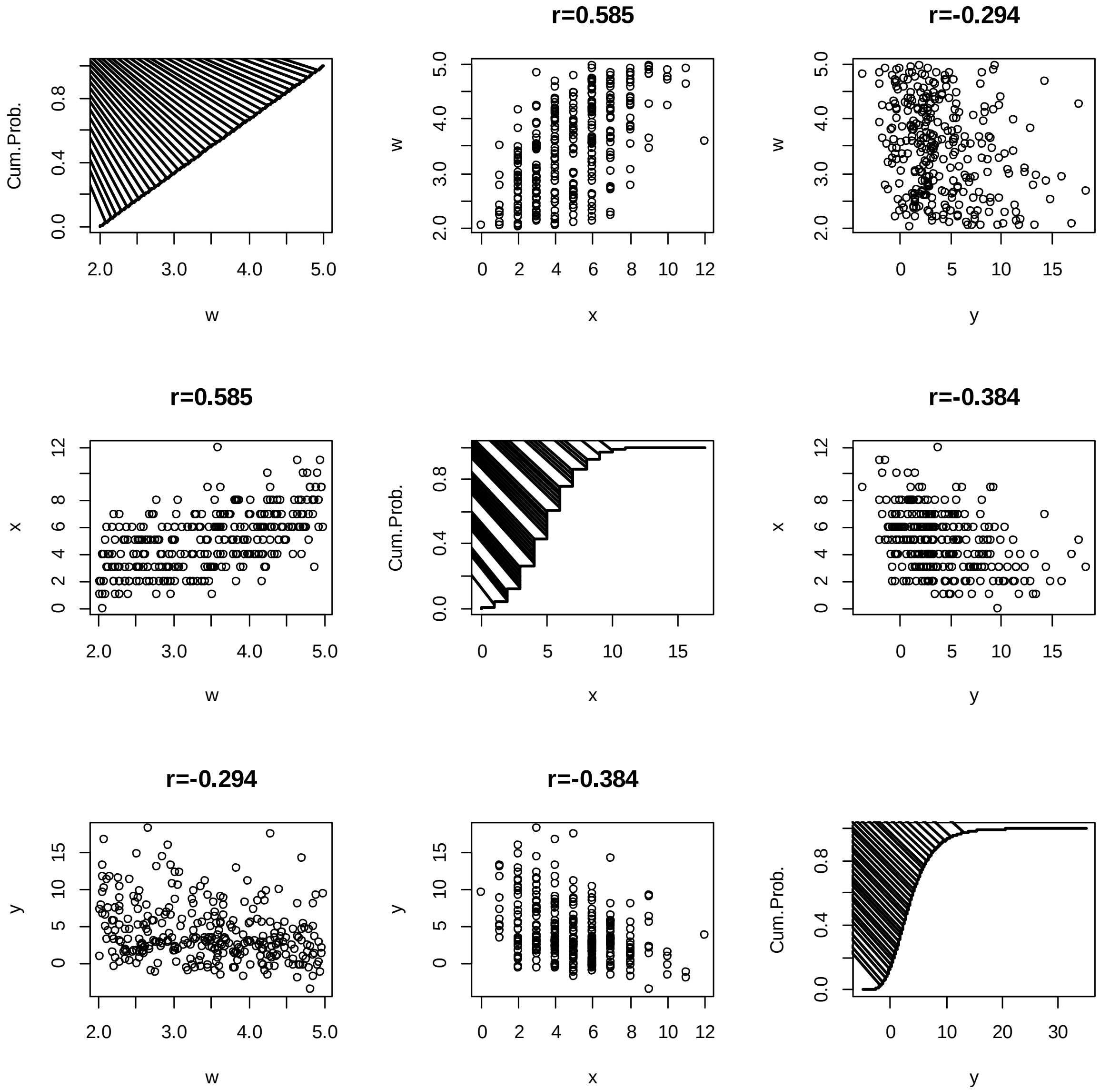
$\text{corr}(Z_1, Z_2) = r$

SKIP

Modeling dependencies

- Need to check that correlations are not infeasible (e.g., corr(A,B) = 0.9, corr(A,C)=0.9, corr(B,C)=–0.9)
- Monte Carlo simulations can model each of the special cases, but not complex dependence
- Many dependence patterns yield the same correlation coefficient (most MC software arbitrarily selects one of these patterns)
- It can be difficult to specify the dependence if empirical information is sparse

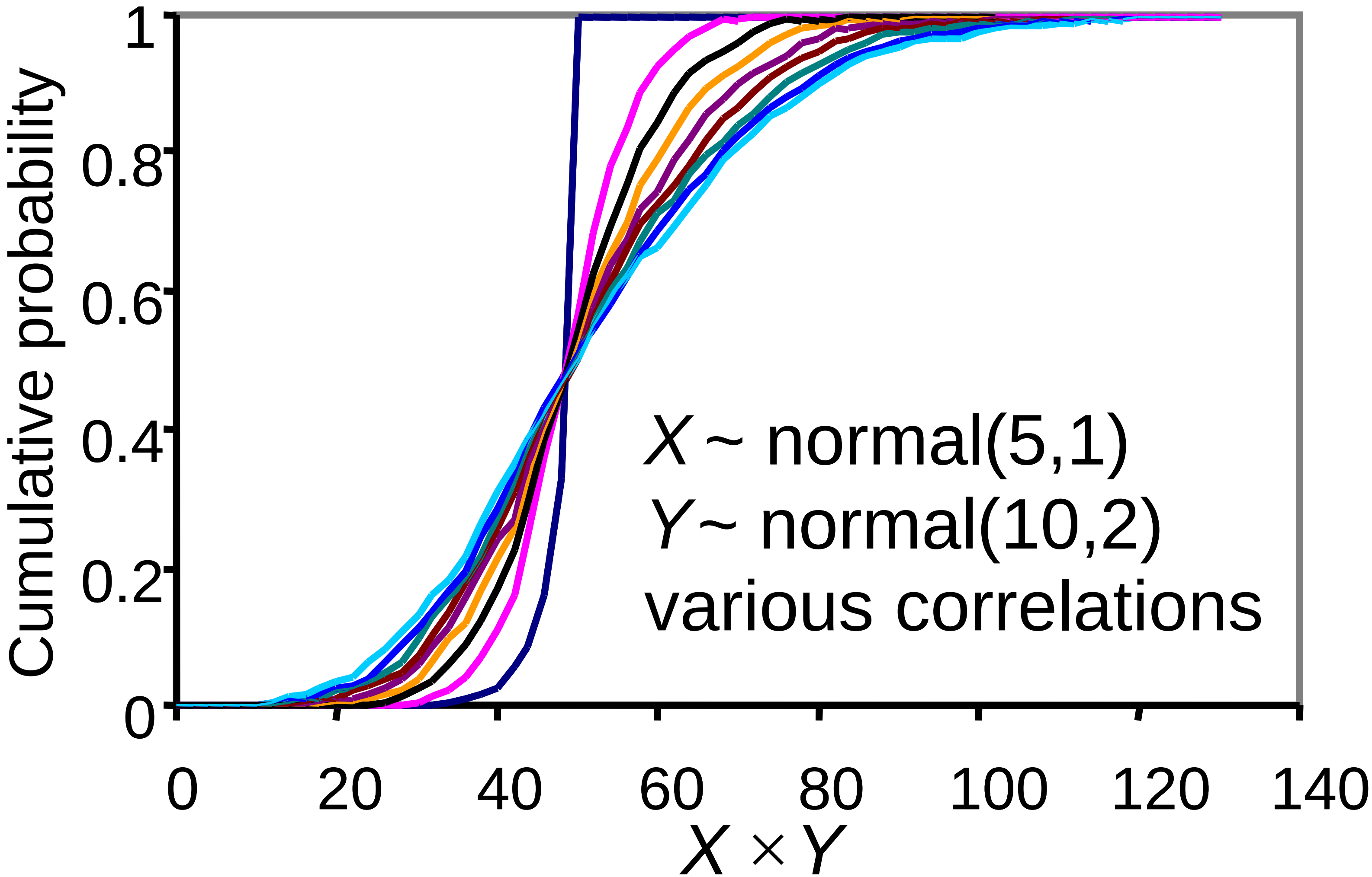
Correlations with sra.r



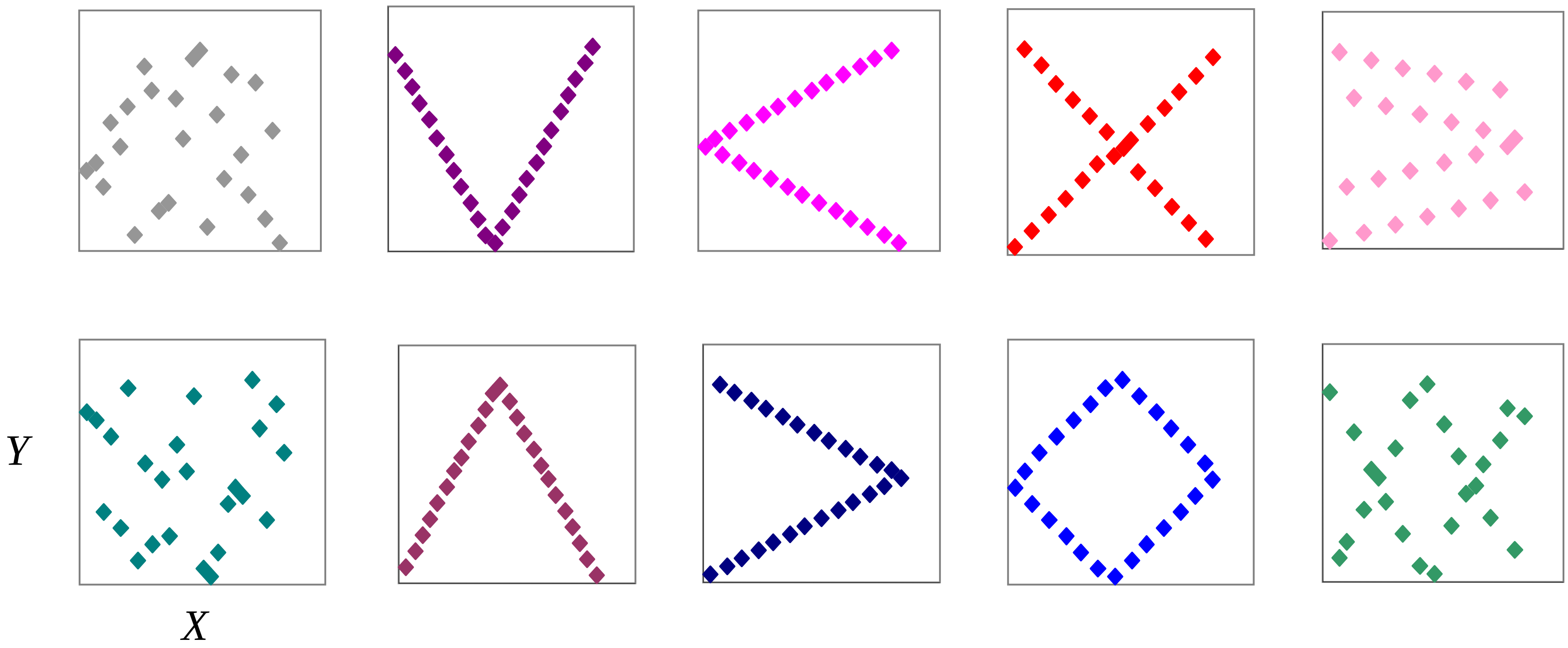
Script to make complex correlations

```
S = c( 1.00,  0.60, -0.30,
            0.60,  1.00, -0.40,
           -0.30, -0.40,  1.00)
correlations(S)
w = uniform(2,5, r=MC$r[,1])
x = poisson(5,   r=MC$r[,2])
y = gumbel(2,3,  r=MC$r[,3])
plotcorrs(c(w,x,y),names=c('w','x','y'))
```

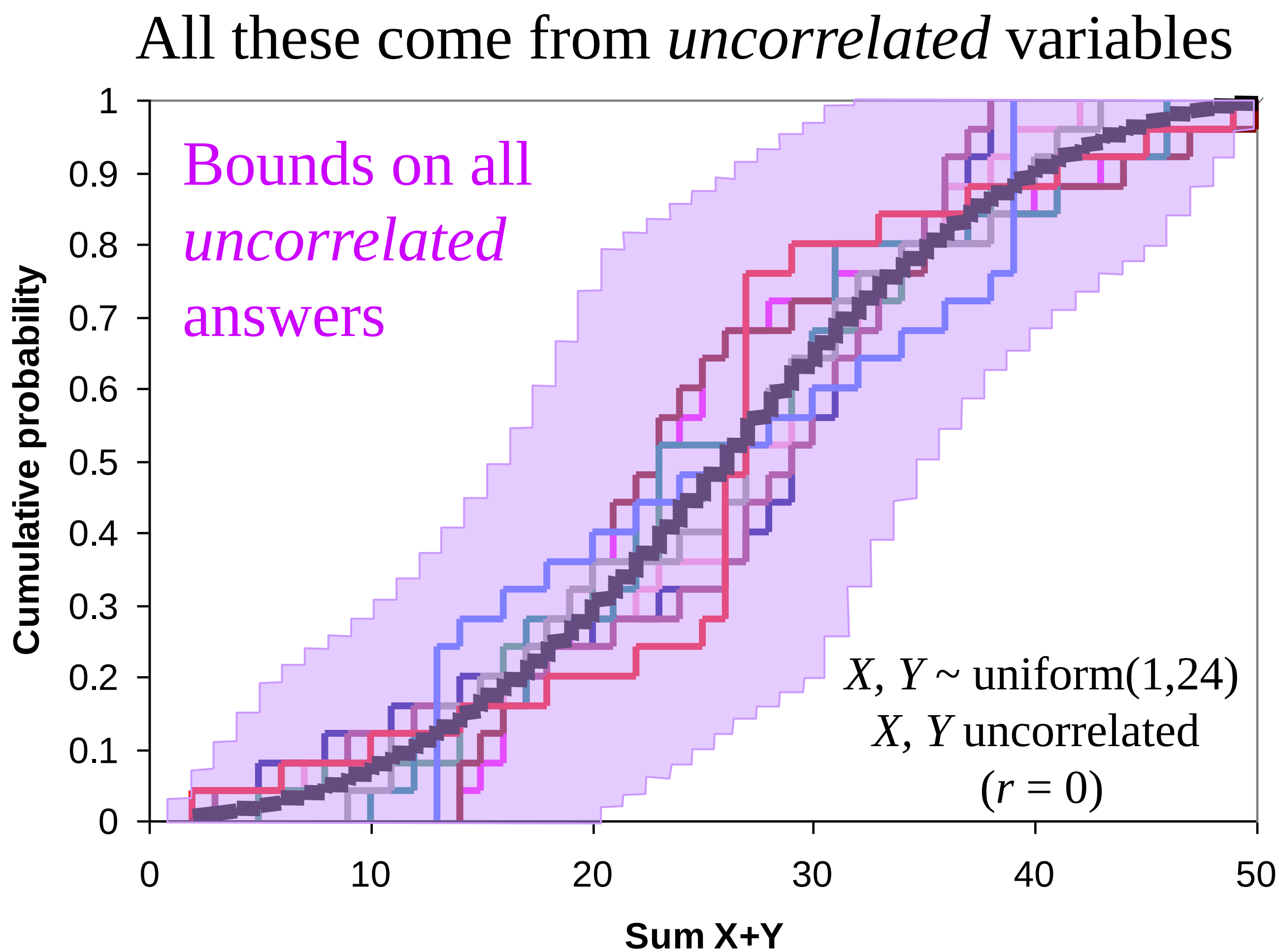
But what if we don't know *r* ?



Scattergrams with *zero* correlation



$X+Y$ depends on dependence



Can’t always assume independence

- If you know the mechanism of dependence, you can model it
- Else, have to reproduce the statistical patterns
- If dependence unknown, can’t use either way
- Even small correlations can have a big effect on convolutions

What about other dependencies?

- Independent
- Perfectly positive (comonotonic)
- Opposite (countermonotonic)
- Positively or negatively associated
- Specified correlation coefficient
- Nonlinear dependence (copula)
- Unknown dependence

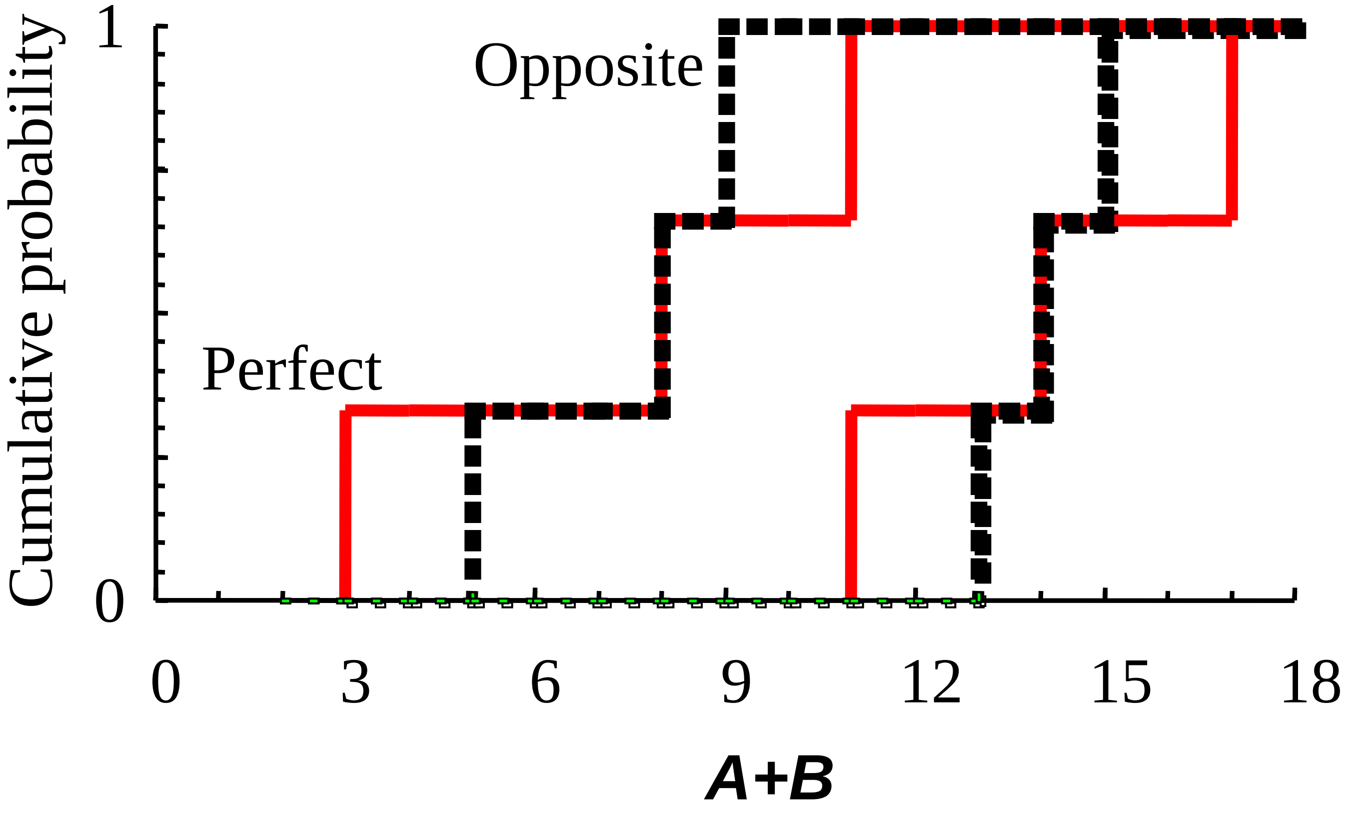
Perfect dependence

$A+B$ perfect positive	$A\in[1,3]$ $p_1 = 1/3$	$A\in[2,4]$ $p_2 = 1/3$	$A\in[3,5]$ $p_3 = 1/3$
$B\in[2,8]$ $q_1 = 1/3$	$A+B\in[3,11]$ prob=1/3	$A+B\in[4,12]$ prob=0	$A+B\in[5,13]$ prob=0
$B\in[6,10]$ $q_2 = 1/3$	$A+B\in[7,13]$ prob=0	$A+B\in[8,14]$ prob=1/3	$A+B\in[9,15]$ prob=0
$B\in[8,12]$ $q_3 = 1/3$	$A+B\in[9,15]$ prob=0	$A+B\in[10,16]$ prob=0	$A+B\in[11,17]$ prob=1/3

Opposite dependence

$A+B$ opposite positive	$A\in[1,3]$ $p_1 = 1/3$	$A\in[2,4]$ $p_2 = 1/3$	$A\in[3,5]$ $p_3 = 1/3$
$B\in[2,8]$ $q_1 = 1/3$	$A+B\in[3,11]$ prob=0	$A+B\in[4,12]$ prob=0	$A+B\in[5,13]$ prob=1/3
$B\in[6,10]$ $q_2 = 1/3$	$A+B\in[7,13]$ prob=0	$A+B\in[8,14]$ prob=1/3	$A+B\in[9,15]$ prob=0
$B\in[8,12]$ $q_3 = 1/3$	$A+B\in[9,15]$ prob= 1/3	$A+B\in[10,16]$ prob=0	$A+B\in[11,17]$ prob=0

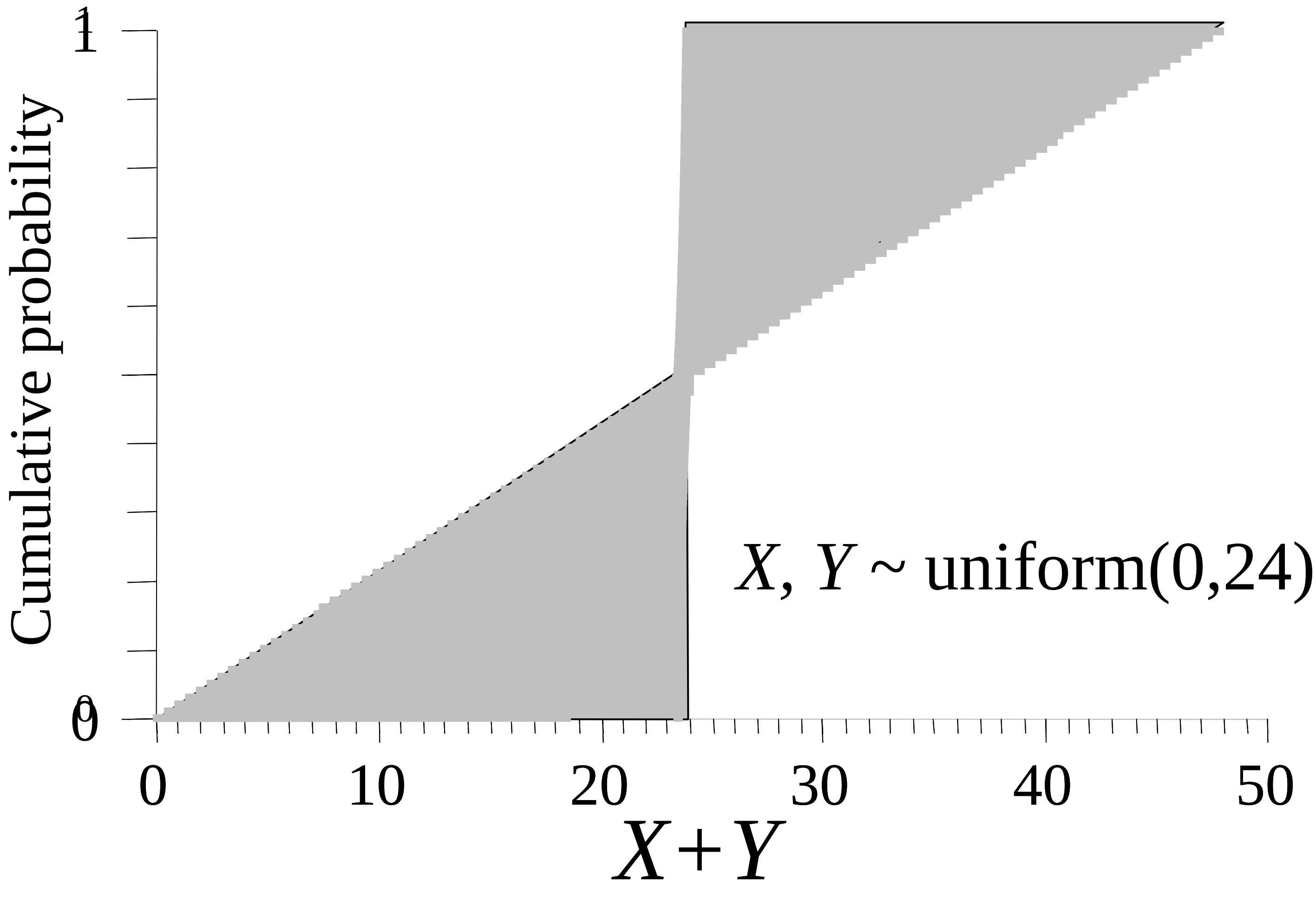
Perfect and opposite dependencies



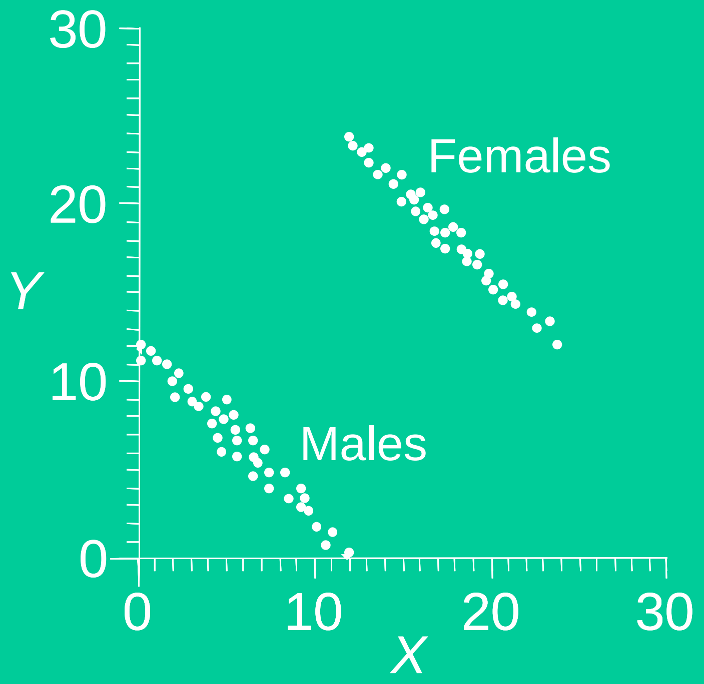
Uncertainty about dependence

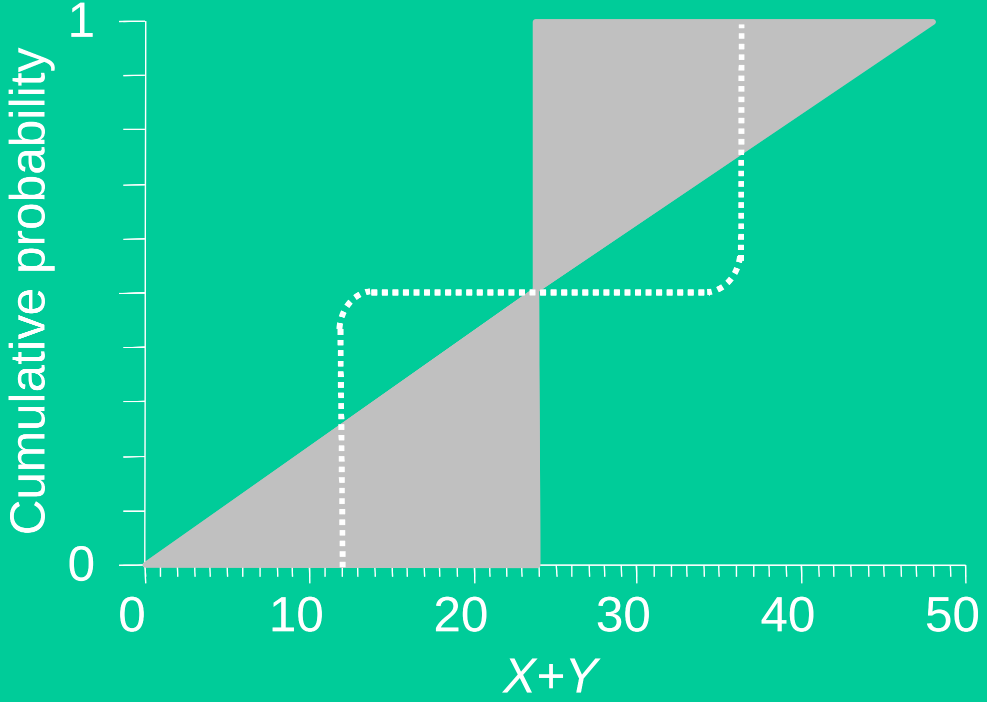
- Sensitivity analyses usually used
  - Vary correlation coefficient between –1 and +1
- But this *underestimates* the true uncertainty
  - Example: suppose X, $Y \sim$ uniform(0,24) but we don't know the dependence between X and Y

Varying the correlation coefficient



Counterexample: outside the cone!



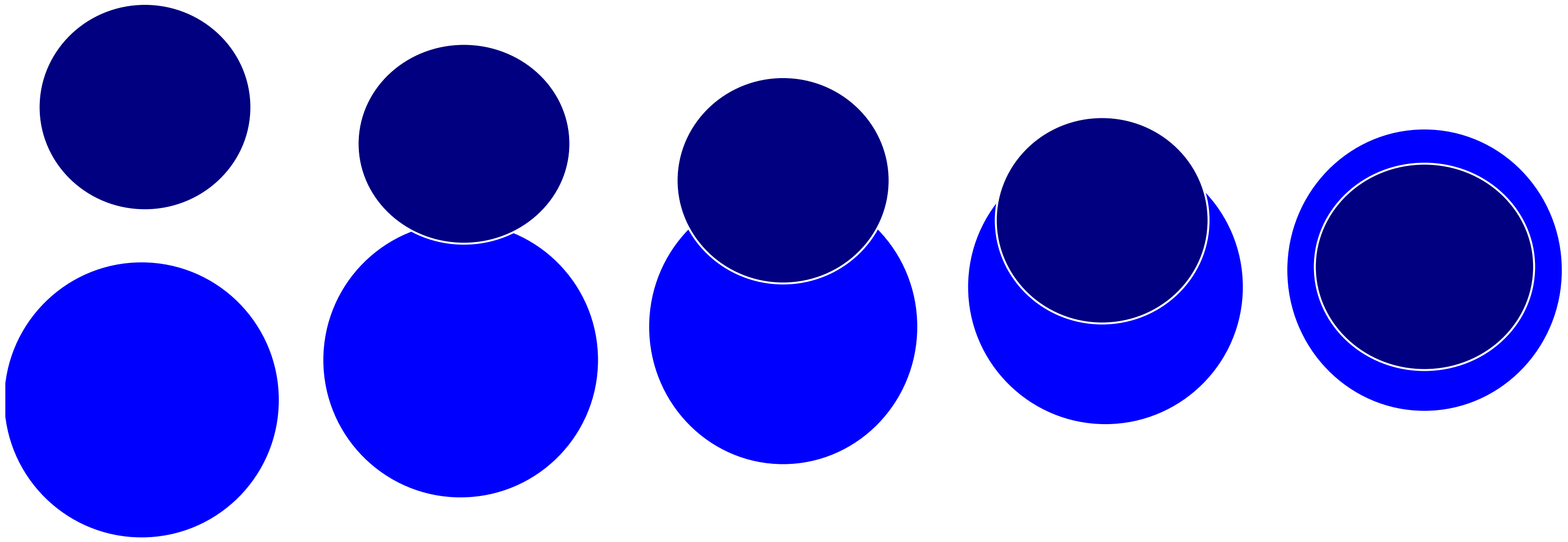


Fréchet inequalities

They make *no assumption* about dependence (Fréchet 1935)

$$\max(0,\ \mathrm{P}(A)+\mathrm{P}(B)-1) \le \mathrm{P}(A\ \&\ B) \le \min(\mathrm{P}(A),\ \mathrm{P}(B))$$

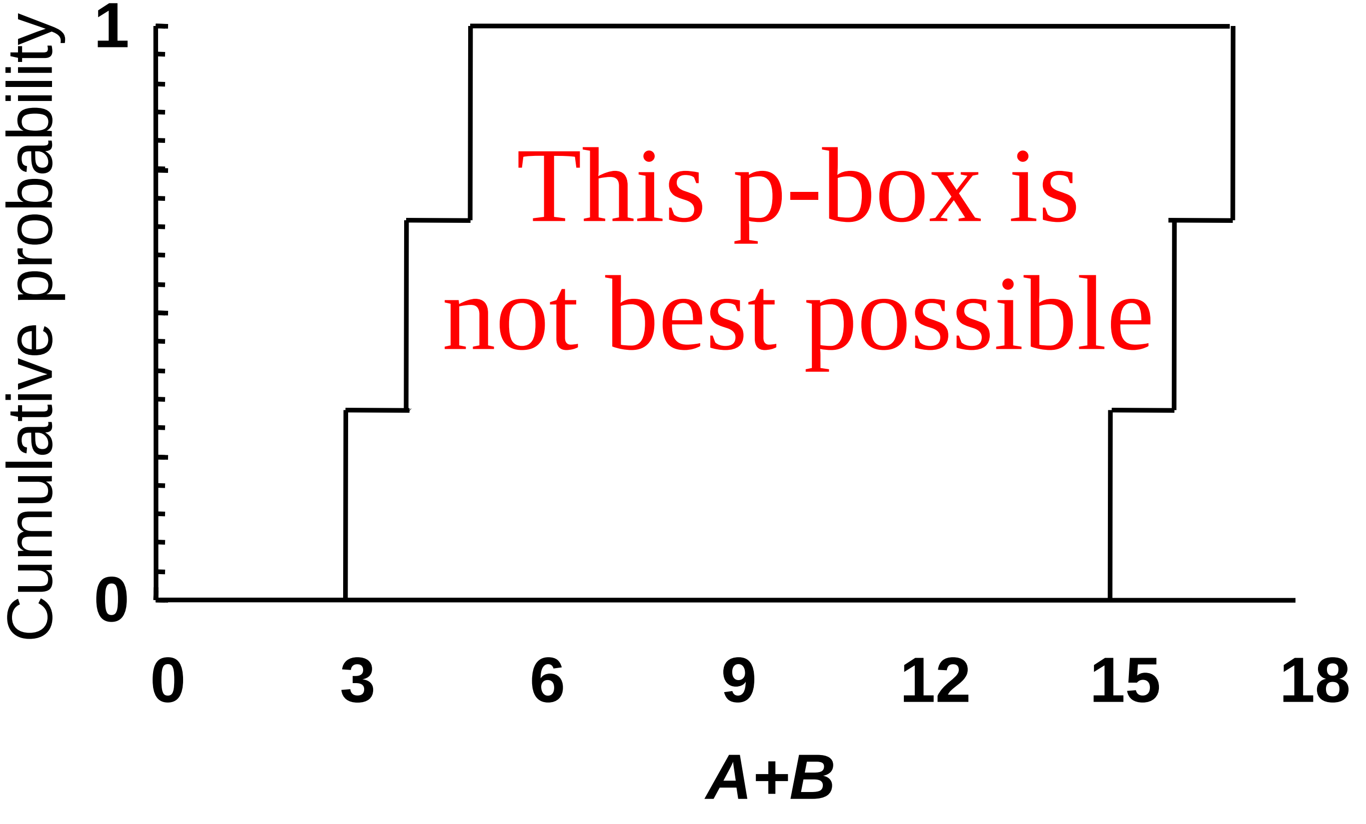
$$\max(\mathrm{P}(A),\ \mathrm{P}(B)) \le \mathrm{P}(A \vee B) \le \min(1,\ \mathrm{P}(A)+\mathrm{P}(B))$$



Fréchet case (no assumption)

$A+B$ Fréchet case	$A\in[1,3]$ $p_1 = 1/3$	$A\in[2,4]$ $p_2 = 1/3$	$A\in[3,5]$ $p_3 = 1/3$
$B\in[2,8]$ $q_1 = 1/3$	$A+B\in[3,11]$ prob=[0,1/3]	$A+B\in[4,12]$ prob=[0,1/3]	$A+B\in[5,13]$ prob=[0,1/3]
$B\in[6,10]$ $q_2 = 1/3$	$A+B\in[7,13]$ prob=[0,1/3]	$A+B\in[8,14]$ prob=[0,1/3]	$A+B\in[9,15]$ prob=[0,1/3]
$B\in[8,12]$ $q_3 = 1/3$	$A+B\in[9,15]$ prob=[0,1/3]	$A+B\in[10,16]$ prob=[0,1/3]	$A+B\in[11,17]$ prob=[0,1/3]

Naïve Fréchet case



Fréchet can be improved

- Interval estimates of probabilities don’t reflect the fact that the sum must equal one
- Resulting p-box is too fat
- Linear programming needed to get the optimal answer using this approach

- Frank, Nelsen and Sklar (1987) gave a way to compute the optimal answer directly

Yikes!

SKIP

Frank, Nelsen and Sklar (1987)

Suppose $X \sim F$ and $Y \sim G$.
If X and Y are independent, then the distribution of $X+Y$ is

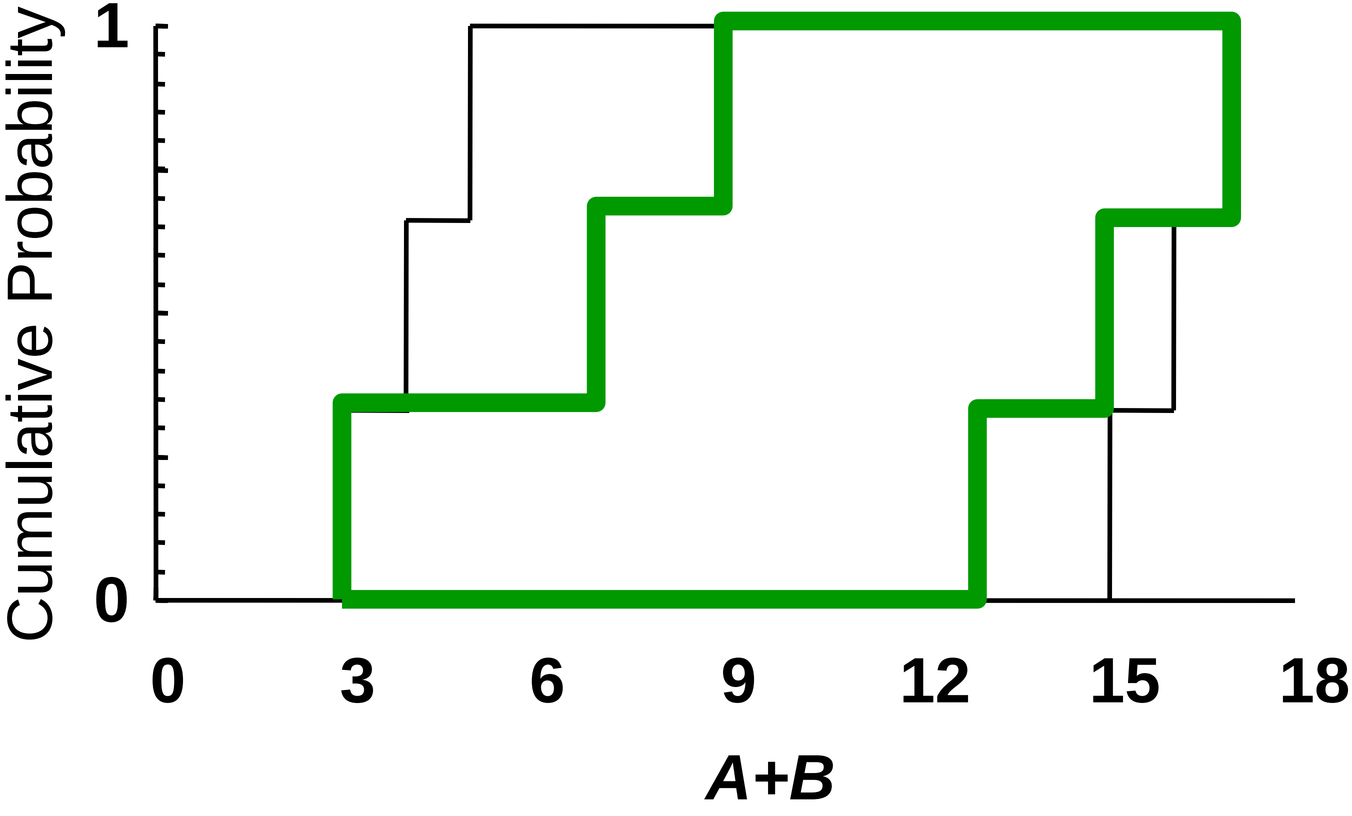
$$\sigma_{+,C}(F,G)(z) = \int_{x+y<z} dC\big(F(x),G(y)\big)$$

In any case, and irrespective of their dependence, this distribution is bounded by

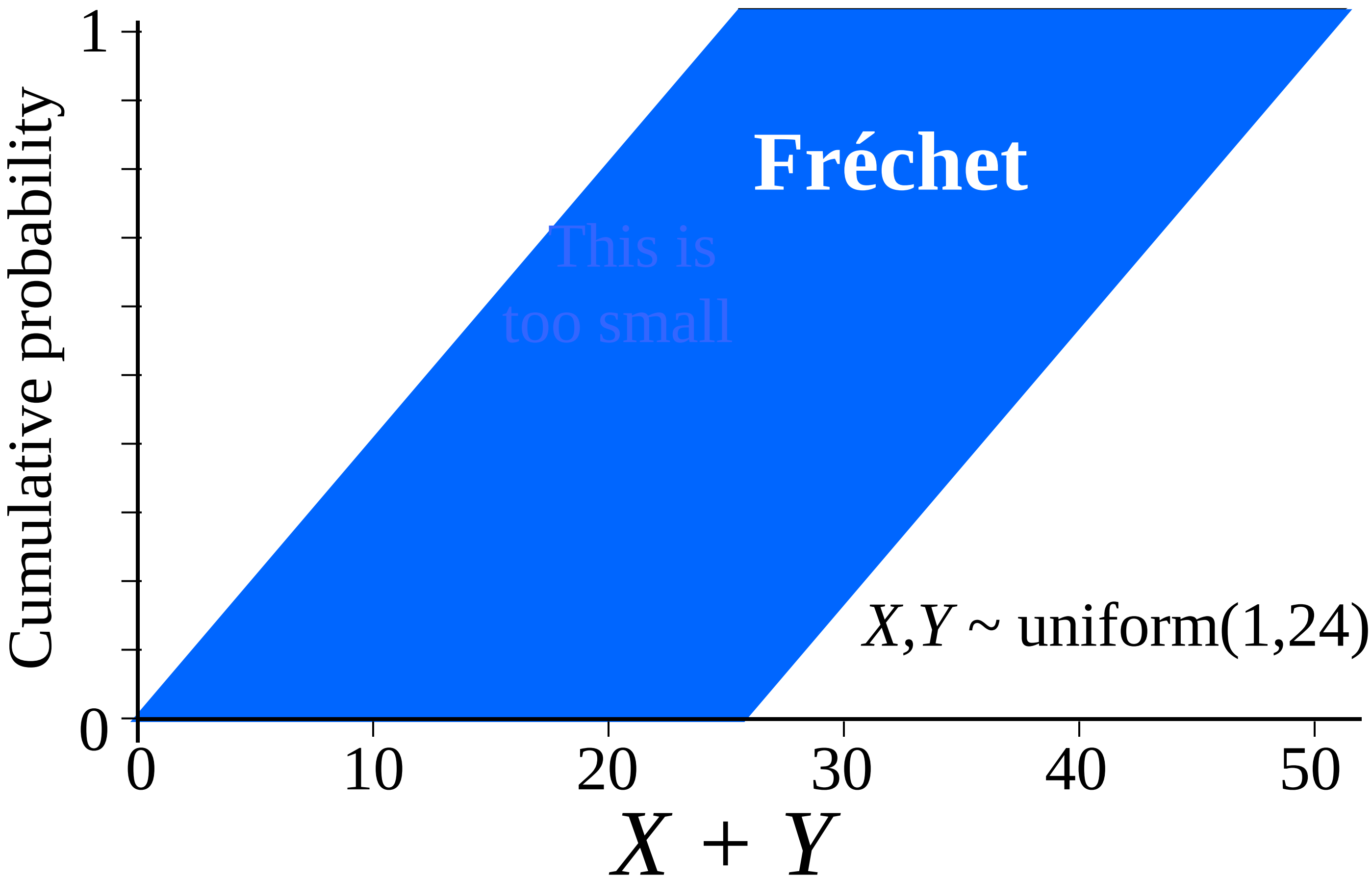
$$\left[\sup_{z=x+y} \max\big(F(x)+G(y)-1,\,0\big),\ \inf_{z=x+y} \min\big(F(x)+G(y),1\big)\right]$$

This formula can be generalized to work with bounds on F and G.

Best possible bounds



Unknown dependence



Fréchet dependence bounds

- Cannot be obtained by sensitivity studies
- Guaranteed to enclose results no matter what correlation or dependence there may be between the variables
- Best possible (couldn't be any tighter without saying more about the dependence)
- Can be combined with independence assumptions between other variables

Between independence and Fréchet

- Some information may be available by which the p-boxes could be tightened over the Fréchet case without specifying the dependence perfectly, e.g.,

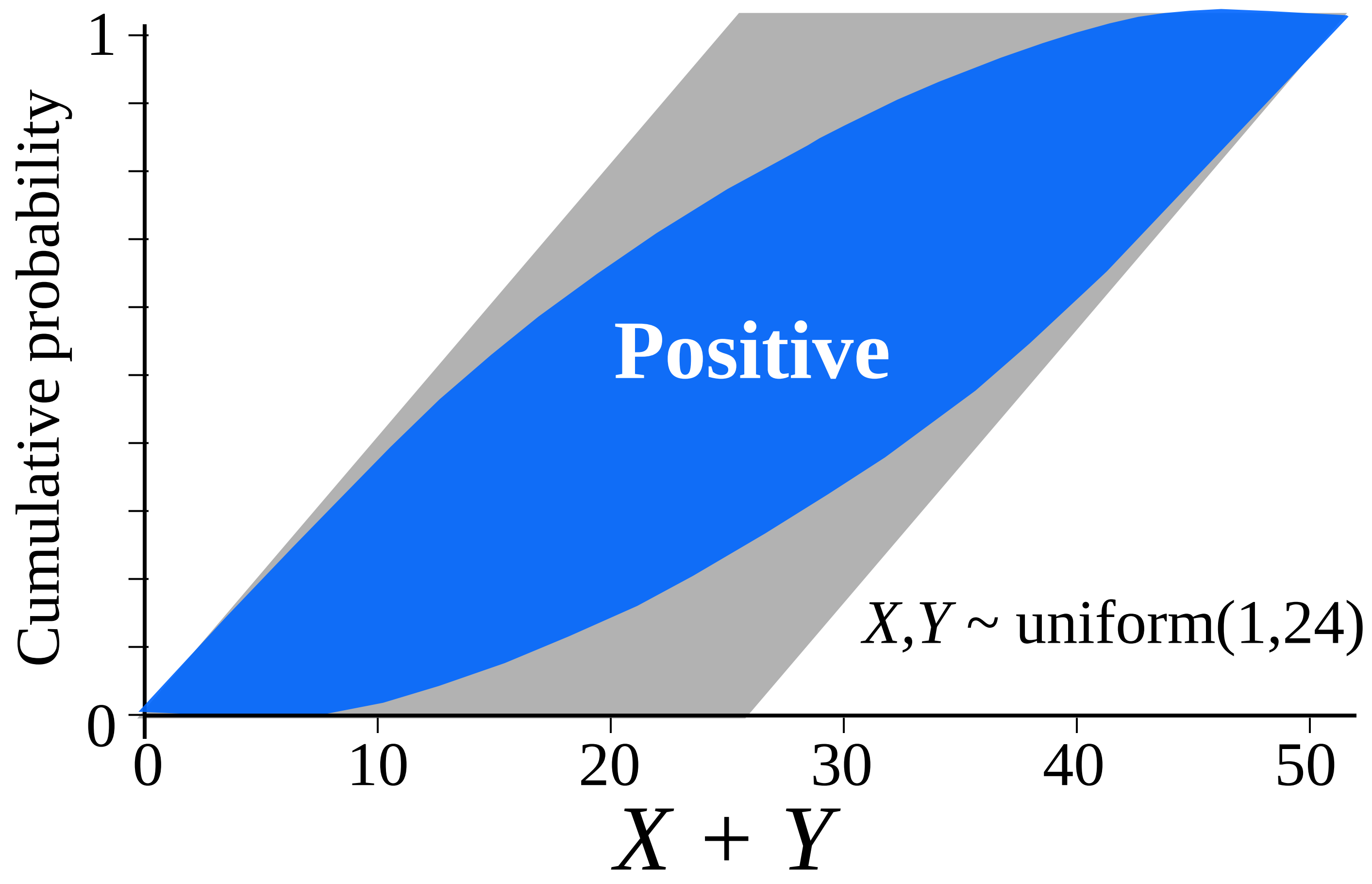
- Dependence is positive (PQD)

  P($X \leq x$, $Y \leq y$) $\geq$ P($X \leq x$) P($Y \leq y$) for all x and y

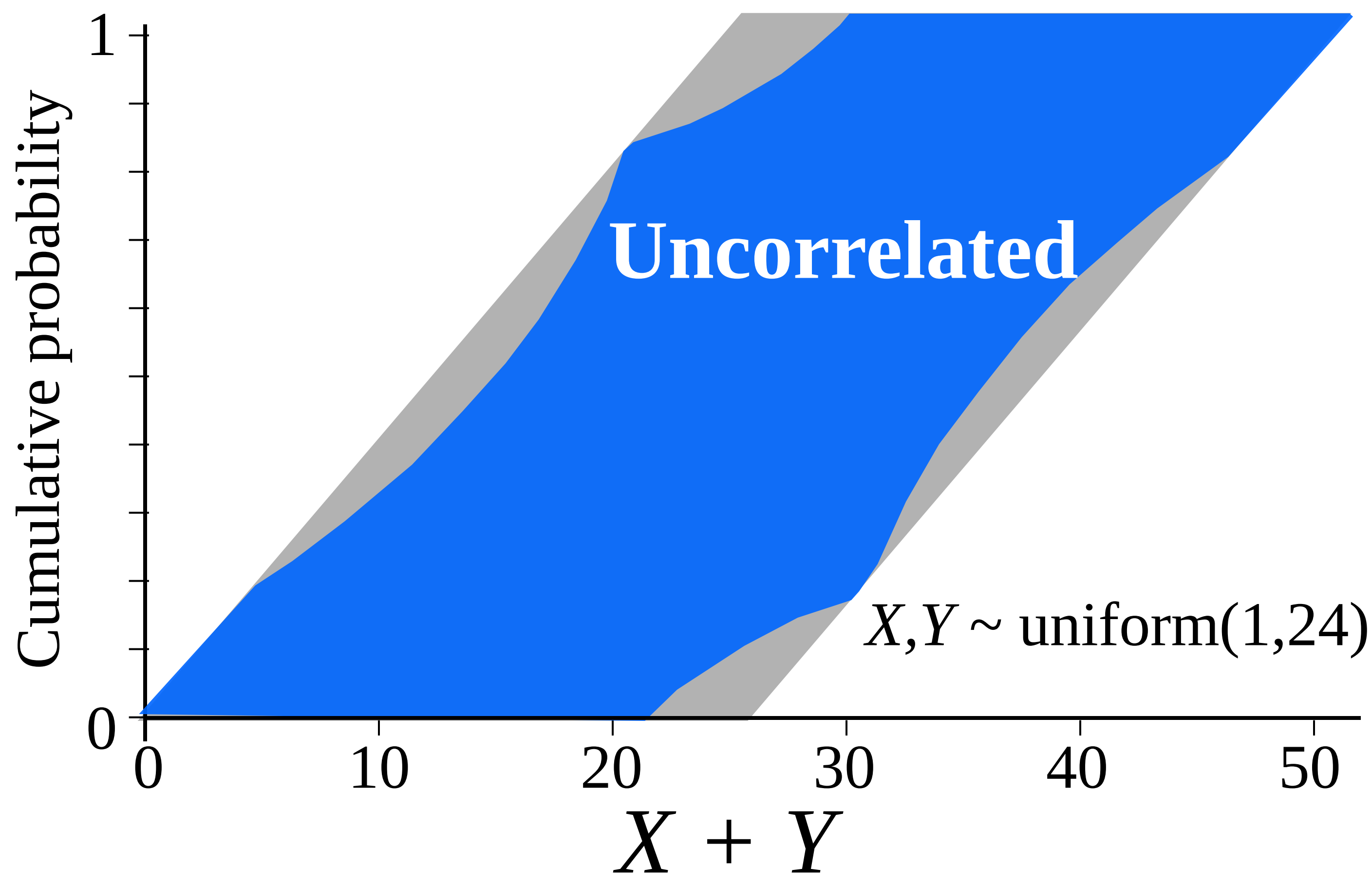
- Variables are uncorrelated

  Pearson correlation r is zero

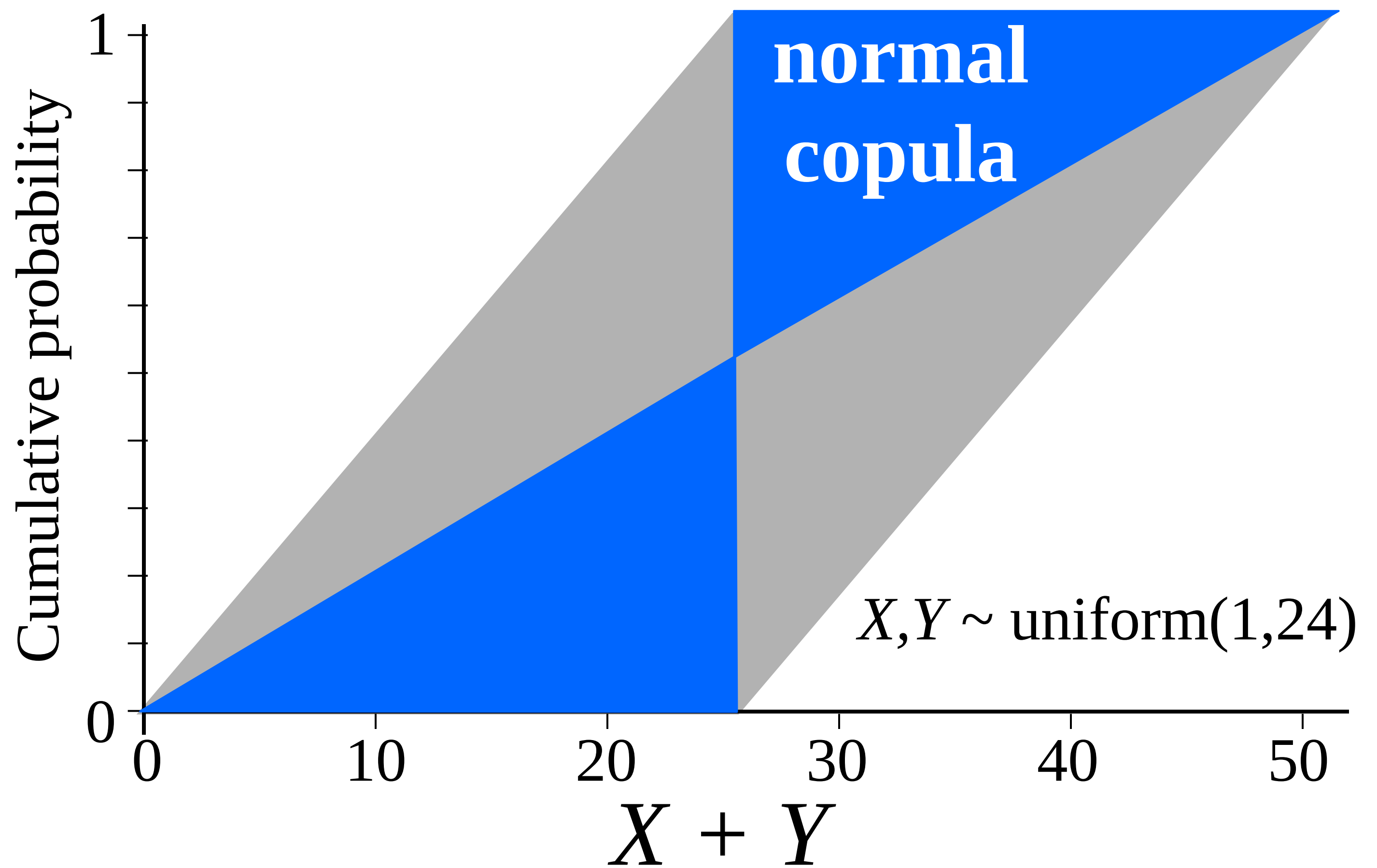
Unknown but positive dependence



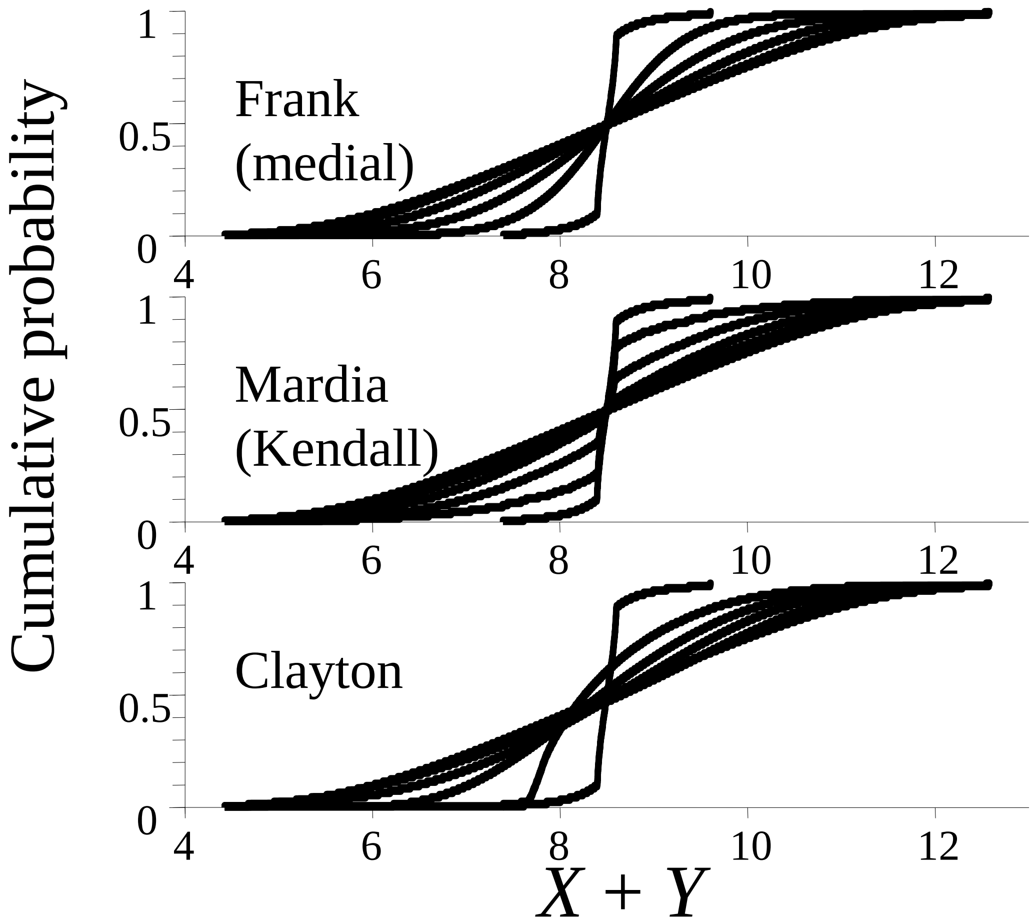
Uncorrelated variables



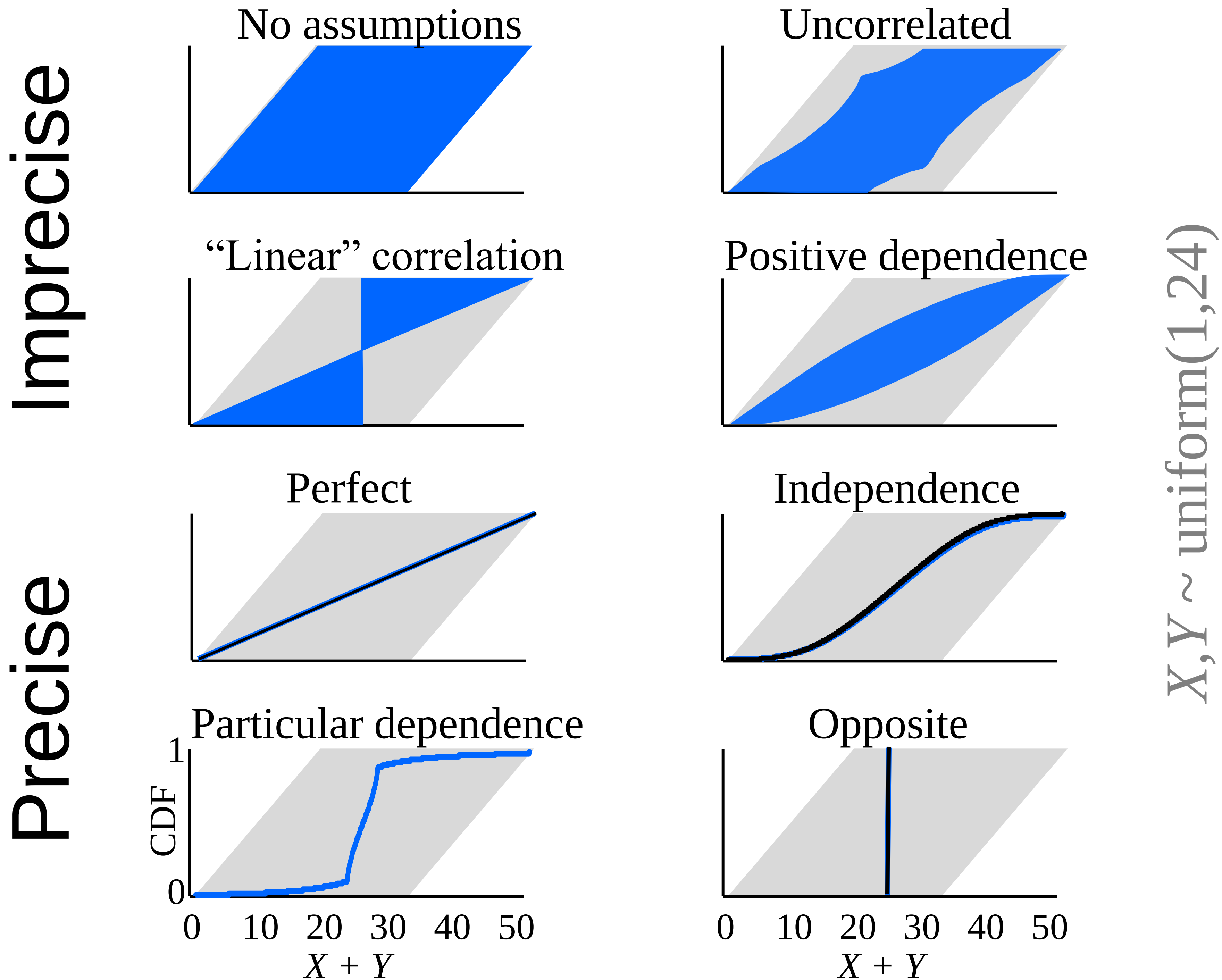
“Linear” correlation



Can model dependence exactly too



X ~ normal(5,1)
Y ~ uniform(2,5)
various correlations
and dependence
functions (copulas)


Imprecise
Precise
No assumptions
Uncorrelated
"Linear" correlation
Positive dependence
Perfect
Independence
Particular dependence
Opposite
CDF
1
0
0
10
20
30
40
50
X + Y
$X, Y \sim$ uniform(1,24)

Example: dioxin inhalation

Location: Superfund site in California

Receptor: adults in neighboring community

Contaminant: dioxin

Exposure route: inhalation of windborne soil

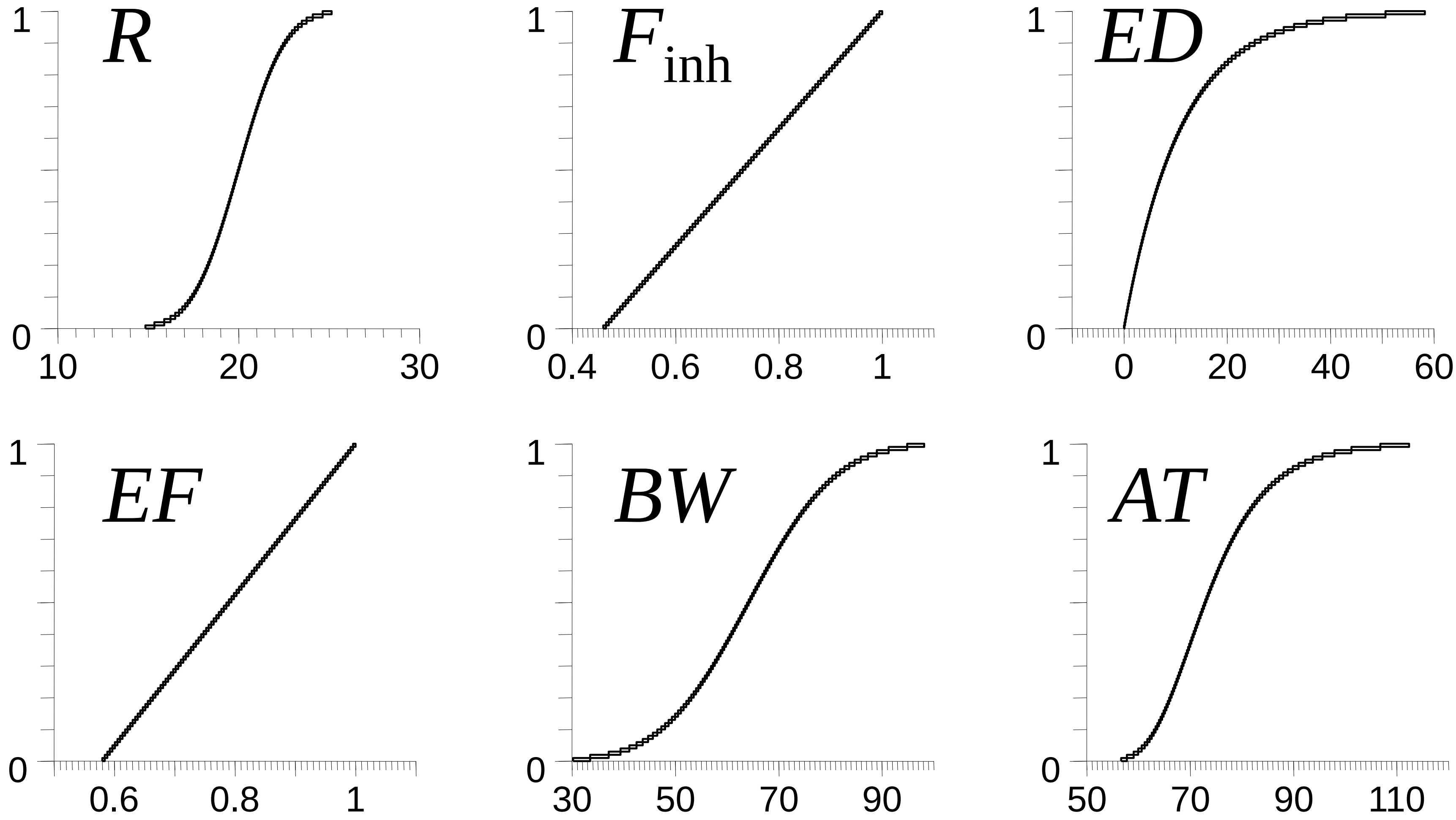
Modified from Table II and IV in Copeland, T.L., A.M. Holbrow, J.M Otani, K.T. Conner and D.J. Paustenbach 1994. Use of probabilistic methods to understand the conservativism in California's approach to assessing health risks posed by air contaminant. *Journal of the Air and Waste Management Association* 44: 1399-1413.

Total daily intake from inhalation

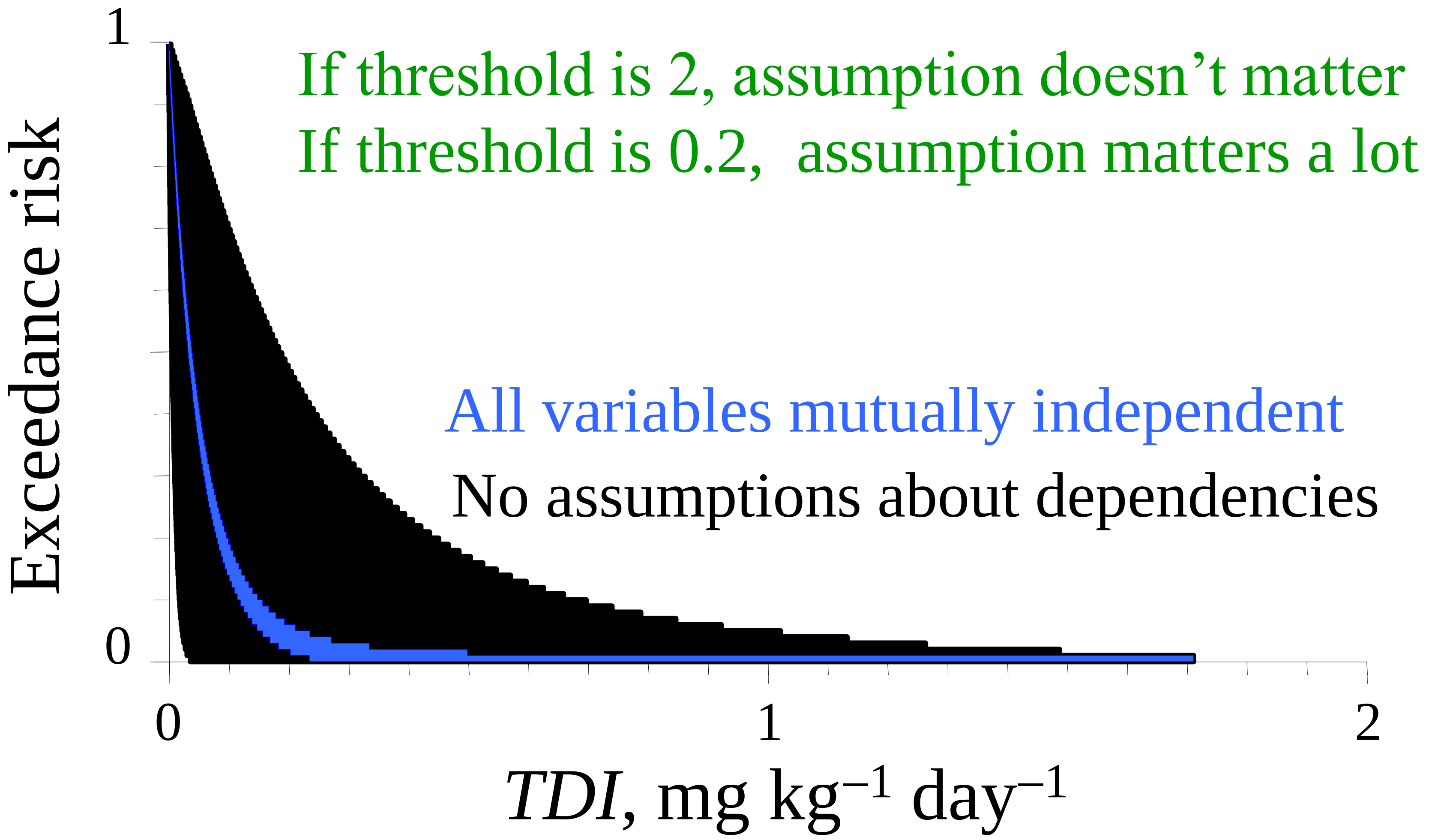
$$TDI = \left(\frac{R \times C_{\mathrm{GL}} \times F_{\mathrm{inh}} \times ED \times EF}{BW \times AT} \right)$$

R = normal(20, 2)	respiration rate, m^3/day
C_{GL} = 2	concentration at ground level, mg/m^3
F_{inh} = uniform(0.46,1)	fraction of particulates retained in lung, [unitless]
ED = exponential(11)	exposure duration, years
EF = uniform(0.58, 1)	exposure frequency, fraction of a year
BW = normal(64.2, 13.19)	receptor body weight, kg
AT = gumbel(70, 8)	averaging time, years

Input distributions



Results


1
0
If threshold is 2, assumption doesn’t matter
If threshold is 0.2, assumption matters a lot
All variables mutually independent
No assumptions about dependencies
Exceedance risk
0
1
2
TDI, mg kg–1 day–1

Uncertainty about dependence

Solved

- Impossible with sensitivity analysis since it’s an infinite-dimensional problem
- Kolmogorov-Fréchet bounding lets you be sure
- Sometimes there’s a big difference, sometimes it’s negligible

What else seems interesting?

- Model uncertainty 19
- Sensitivity analysis 26
- Backcalculation 14
- Interval statistics 28
- Computing with confidence 38
- Fermi estimates 20
- Imprecise probabilities 14
- Nasa case study 19
- EPA case study 29
- Conclusions 15

Model uncertainty

Model uncertainty

- Doubt about the structural form of the model

model uncertainty

- Parameters
- Distribution shape
- Intervariable dependence
- Arithmetic equation
- Level of abstraction

- Usually incertitude rather than variability
- Usually considerable in ecosystems models
- Often the elephant in the middle of the room

Uncertainty in probabilistic analyses

- Parameters
- Data surrogacy
- Distribution shape
- Intervariable dependence
- Arithmetic expression
- Level of abstraction

already

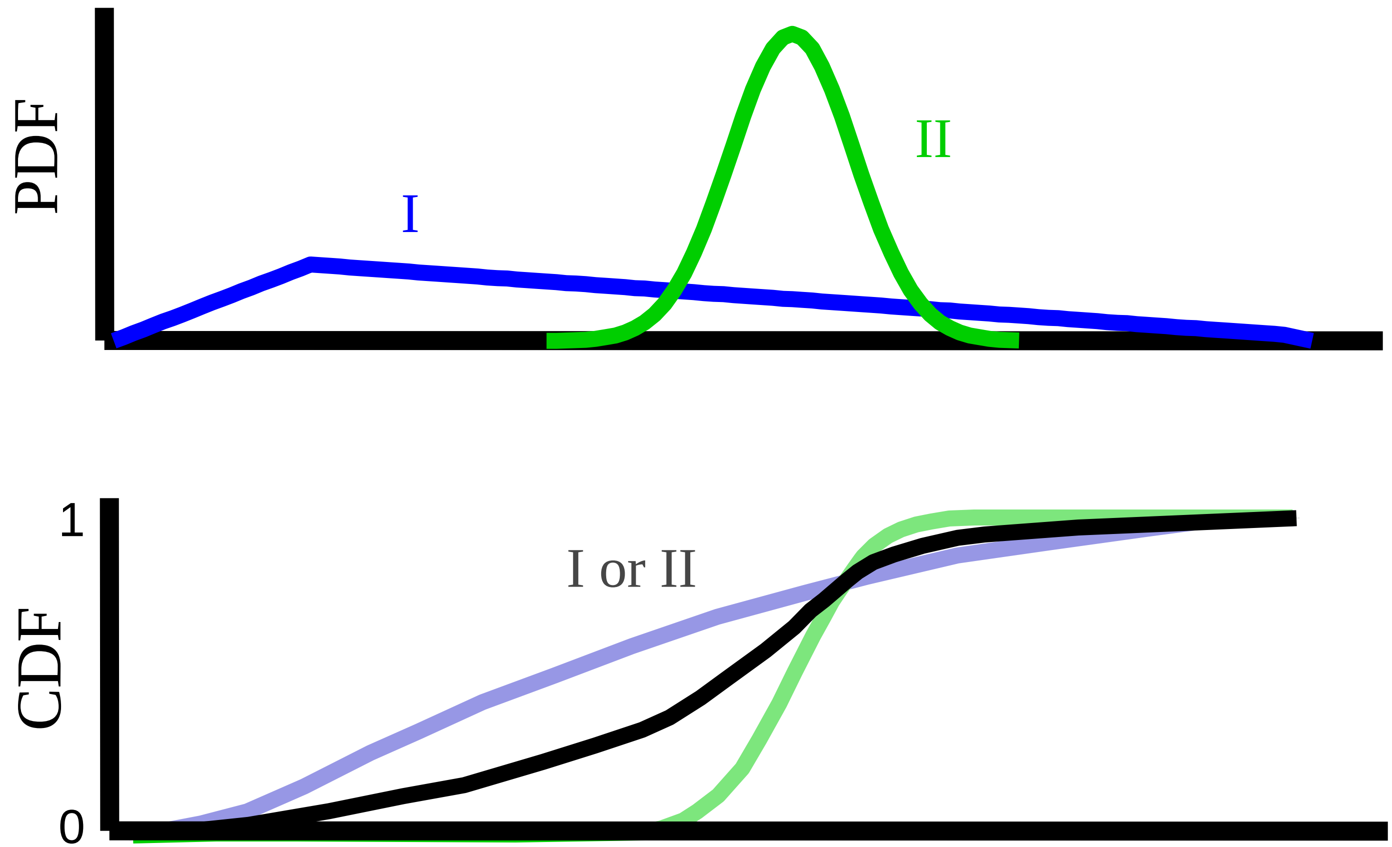
model uncertainty

Monte Carlo strategy

- Introduce a new discrete variable
- Let the value of the variable dictate which model will be used in each rep
- Wiggle the value from rep to rep
- Only works for short, explicit list of models (you have to list the models)
- Many theorists object to this strategy

Model uncertainty as a mixture

If u>0.5 then model=I else model=II



General strategies

- Sensitivity (what-if) studies
- Probabilistic mixture
- Bayesian model averaging
- Enveloping and bounding analyses

Sensitivity (what-if) studies

- Simply re-compute the analysis with alternative assumptions
  - Intergovernmental Panel on Climate Change
- No theory required to use or understand

Drawbacks of what-if

- Consider a long-term model of the economy under global warming stress

  3 baseline weather trends
  3 emission scenarios
  3 population models
  3 mitigation plans

  81 analyses to compute, *and to document*

- Combinatorially complex as more model components are considered

- Cumbersome to summarize results

Probabilistic mixture

NRC uses

- Identify all possible models
- Introduce a new discrete random variable whose value says which model to use; let it vary in MC
- This *averages* probability distributions
- Use weights to account for different credibility (or assume equiprobability)

Drawbacks of mixture

- If you cannot enumerate the possible models, you can’t use this approach
- Averages together incompatible theories and yields an answer that neither theory supports
- Can underestimate tail risks

Bayesian model averaging

- Similar to the probabilistic mixture
- Updates prior probabilities to get weights
- Takes account of available data

Drawbacks of Bayesian averaging

- Requires priors and can be computationally challenging
- Must be able to enumerate the possible models
- Averages together incompatible theories and yields an answer that neither theory supports
- Can underestimate tail risks

Bounding probabilities

- Translate model uncertainties to a choice among distributions
- Envelope the cumulative distributions
- Treat resulting p-box as single object

Drawbacks of bounding

- Cannot account for different model credibilities
- Can’t make use of data
- Doesn’t account for ‘holes’

Numerical example

The function f is one of two possibilities. Either

$$f(A,B) = f_{\text{Plus}}(A,B) = \mathrm{A} + \mathrm{B}$$

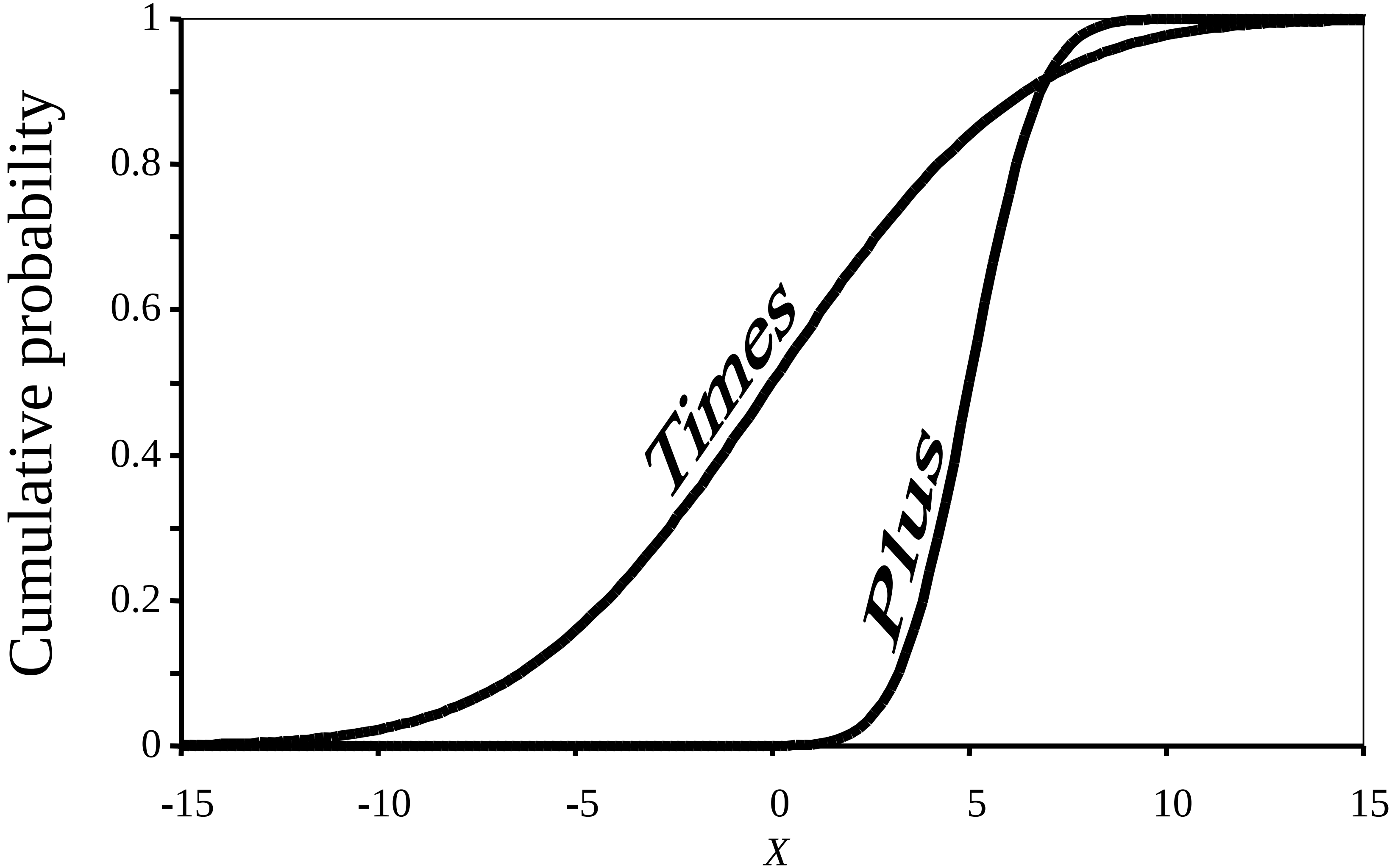
or

$$f(A,B) = f_{\text{Times}}(A,B) = \mathrm{A} \times \mathrm{B}$$

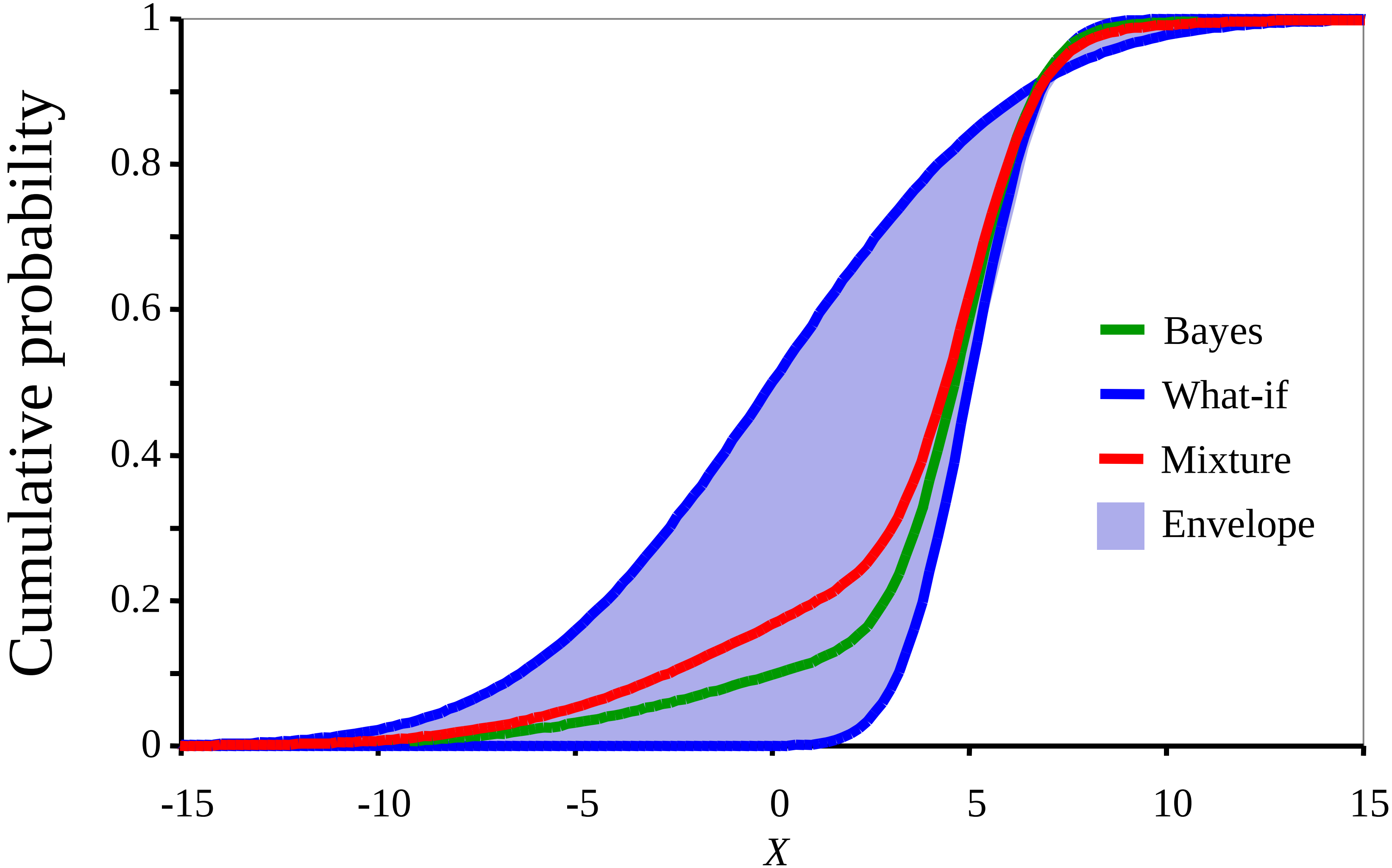
is the correct model, but the analyst does not know which. Suppose that

$$A \sim \text{triangular}(-2.6, 0, 2.6)$$
$$B \sim \text{triangular}(2.4, 5, 7.6).$$

f_{Plus} is twice as likely as f_{Times}; datum: $f(A,B) = 7.59$



Cumulative probability
1
0.8
0.6
0.4
0.2
0
-15
-10
-5
0
5
10
15
X
Times
Plus



Cumulative probability
1
0.8
0.6
0.4
0.2
0
-15
-10
-5
0
5
10
15
X
Bayes
What-if
Mixture
Envelope

When you can enumerate the models

- What-if analysis isn’t feasible in big problems
- Probabilistic mixture is, at best, *ad hoc*
- For abundant data, Bayesian approach is best
- Otherwise, it’s probably just wishful thinking
- Bounding is reliable, but may be too wide

When you can’t list the models

- If you cannot enumerate all the models, bounding is often the only tenable strategy

- Shape of input distributions
- Dependence
- Functional form
  - Laminar versus turbulent flow
  - Linear or nonlinear low-dose extrapolation
  - Ricker versus Beverton-Holt density dependence

Synopsis of the four approaches

- What-if
  - Straightforward, doesn’t conflate uncertainties
  - Must enumerate, combinatorial
- Probabilistic mixture, Bayesian model averaging
  - Single distribution, accounts for data (and priors)
  - Must enumerate, averages incompatible theories
  - Can underestimate tail risks
- Bounding
  - Yields one object; doesn’t conflate or understate risk
  - Cannot account for data or differential credibility

Sensitivity analysis

When to break down and get more data

Justifying further empirical effort

- If the incertitude associated with the results of an analysis is too broad to make practical decisions, and the bounds are best possible, more data is needed
  - *Strong argument for collecting more data*

- Planning empirical efforts can be improved by doing sensitivity analysis of the model

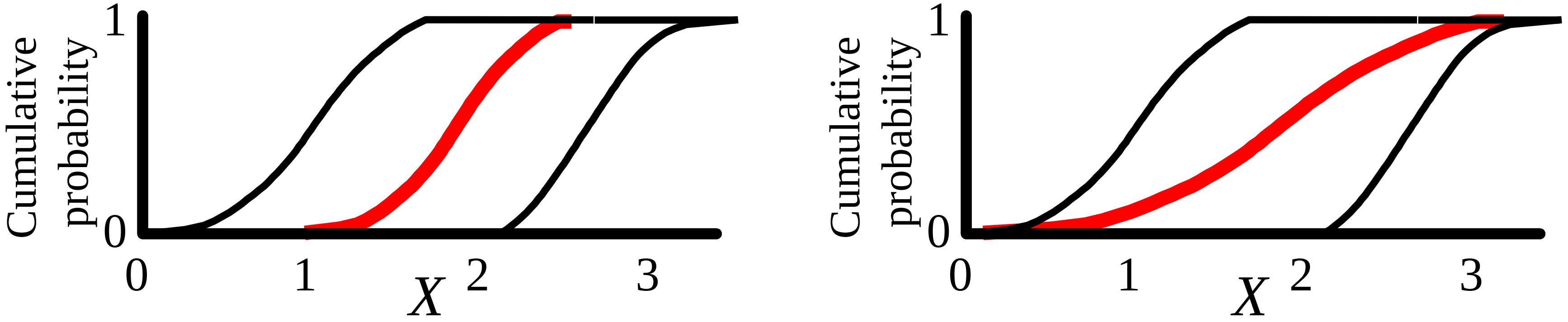
Sensitivity analysis <u>of</u> p-boxes

- Quantifies the reduction in uncertainty of a result when an input is pinched
- *Pinching* is hypothetically replacing it by a less uncertain characterization

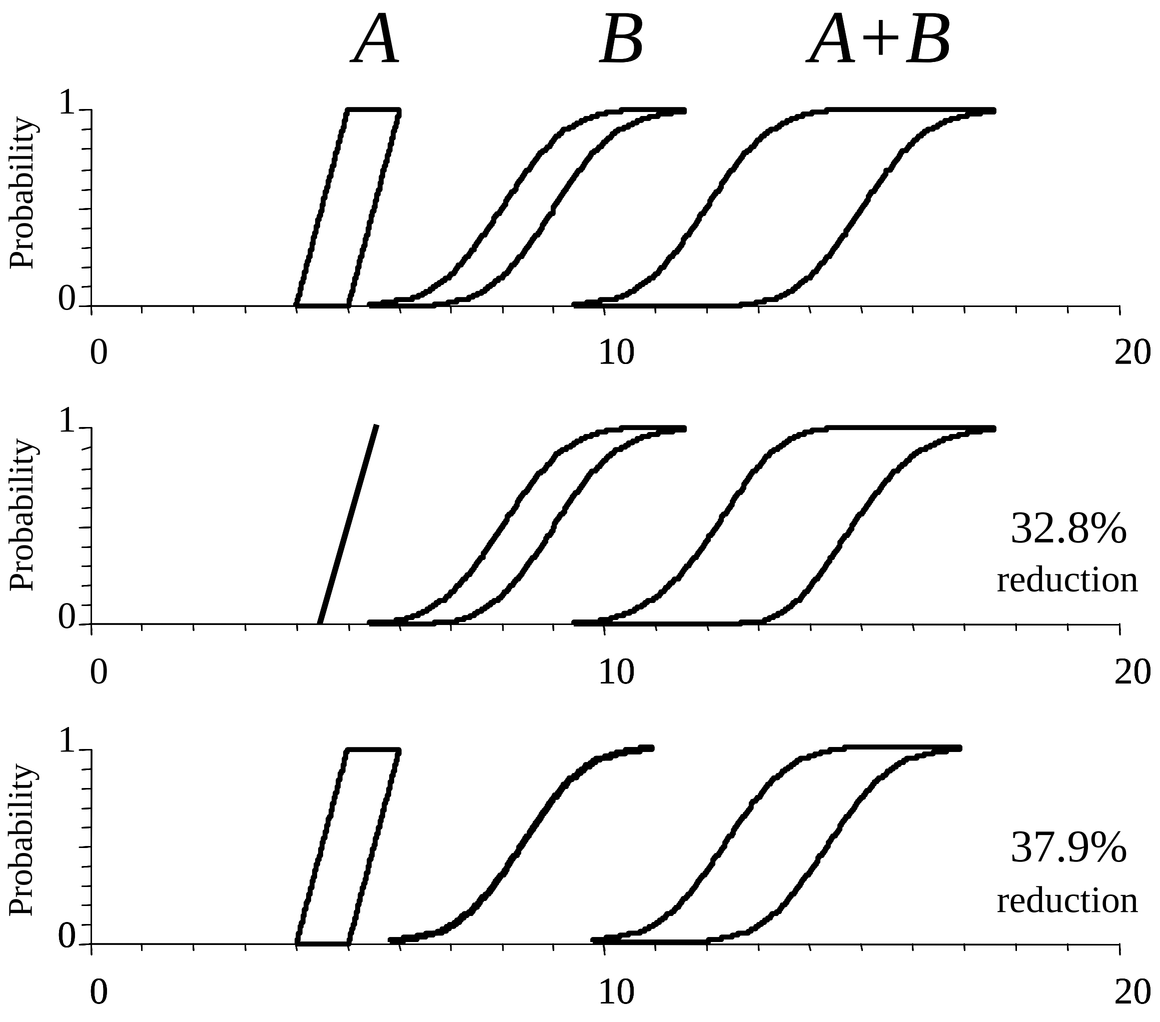
Pinching sensitivity analyses

- Model the possible contraction of incertitude in each input p-box from additional data to be collected
- Recompute analysis with this tighter p-box
  - Others inputs held in their original form
  - Or possibly tighten several if appropriate for planned data
- Estimate improvement in results
  - breadth(tighter input) / breadth(base)
- Repeat for all inputs or data collection strategies
  - Don’t omit variables from sensitivity study or “shortlist”
- Allocate empirical effort by the magnitude of the potential improvement in uncertainty

Pinching to a precise distribution

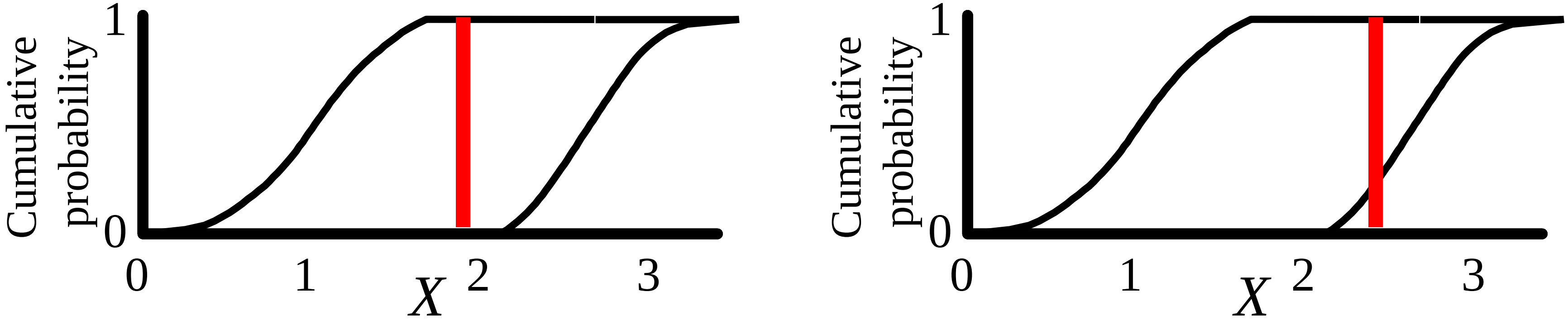


Examples

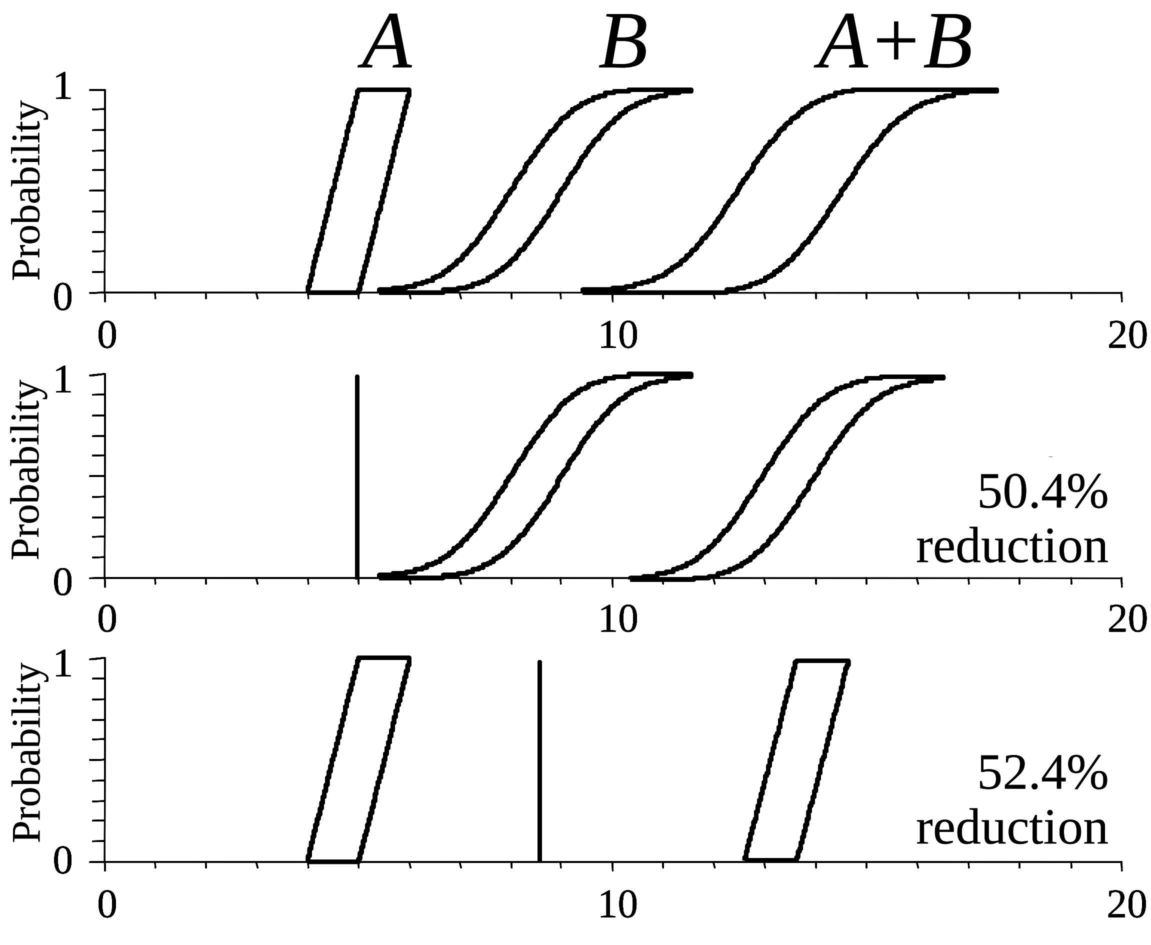


Uses no assumption about dependence between *A* and *B*

Pinching to a point value



Examples



Under independence

Pinching dependence

- An uncertain dependence can also be pinched to a more specific dependence

  positive → independence

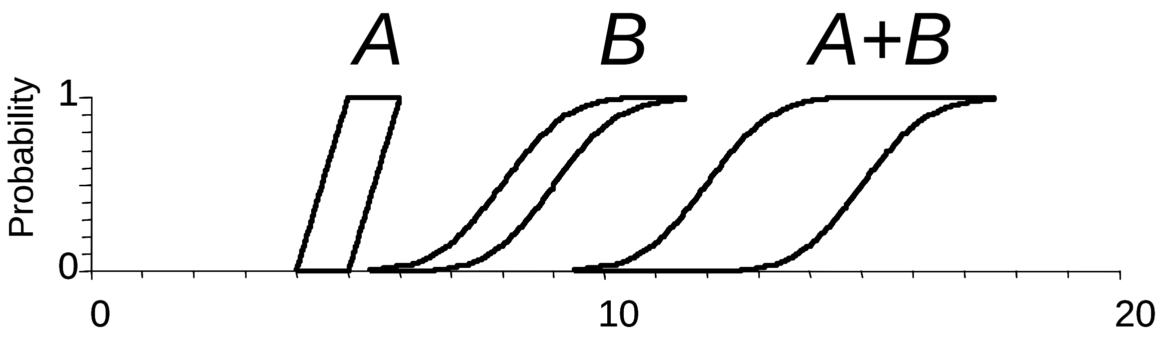
  positive → perfect

  positive → normal copula with correlation 0.3

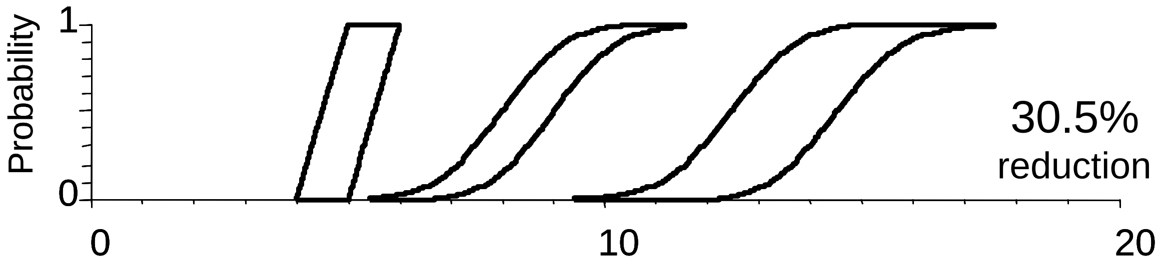
  Fréchet → independence

  etc.

Example

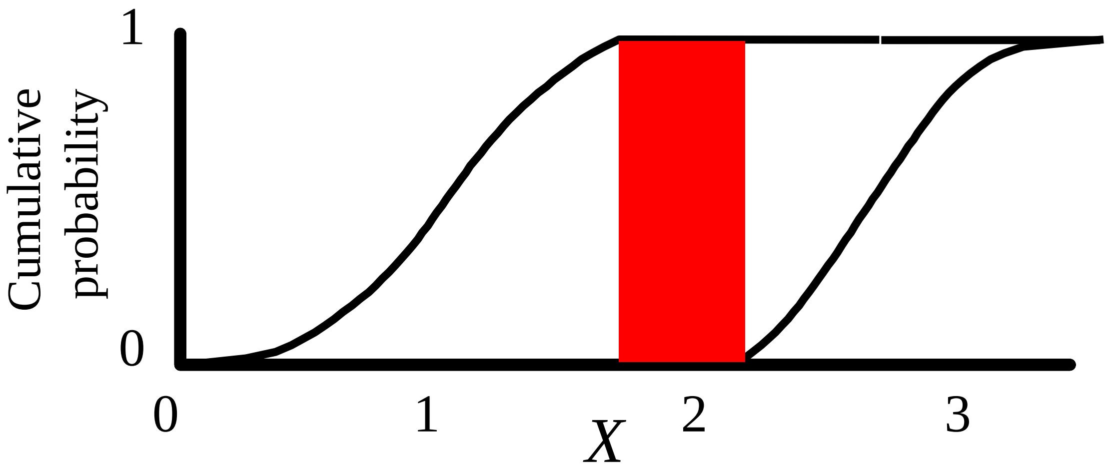


Uses no assumption about dependence between *A* and *B*



Pinch the dependence to *independence*

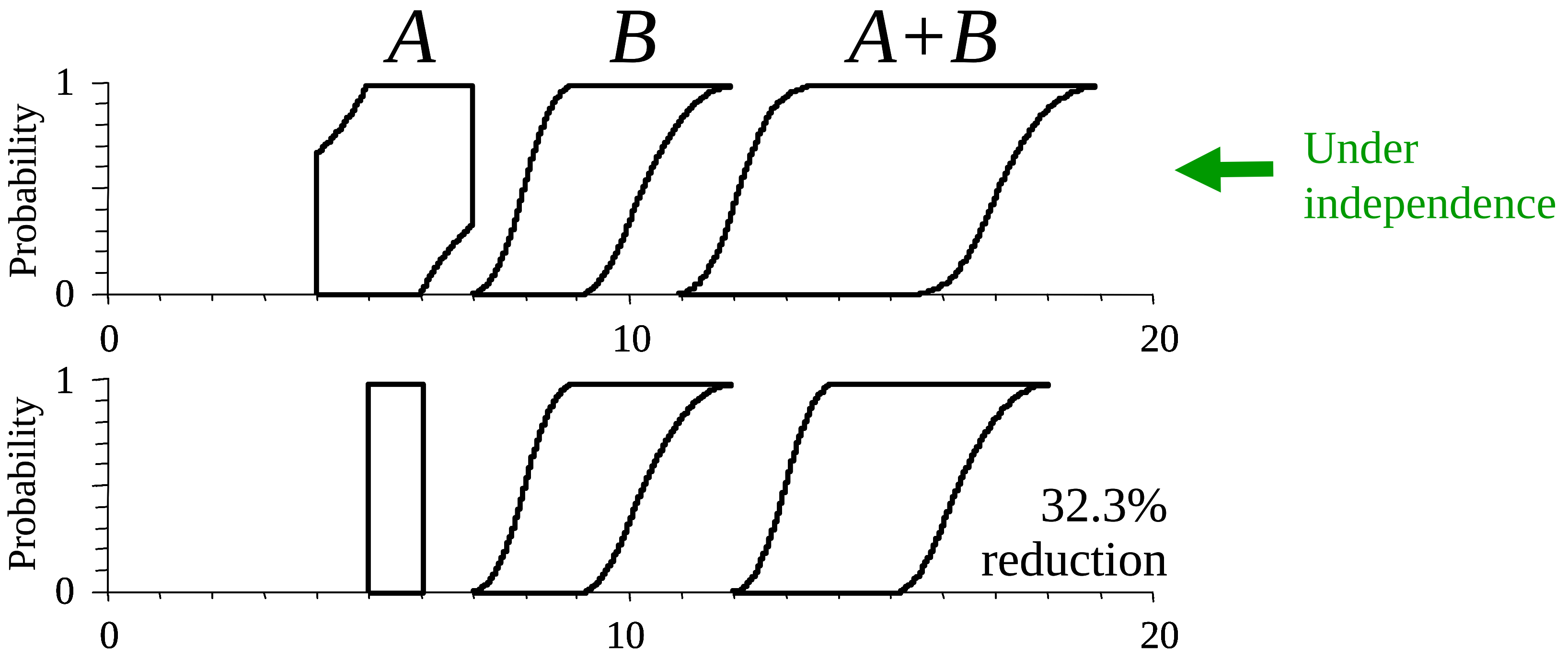
Pinching to a zero-variance interval



Assumes value is constant, but unknown

There's no analog of this in Monte Carlo

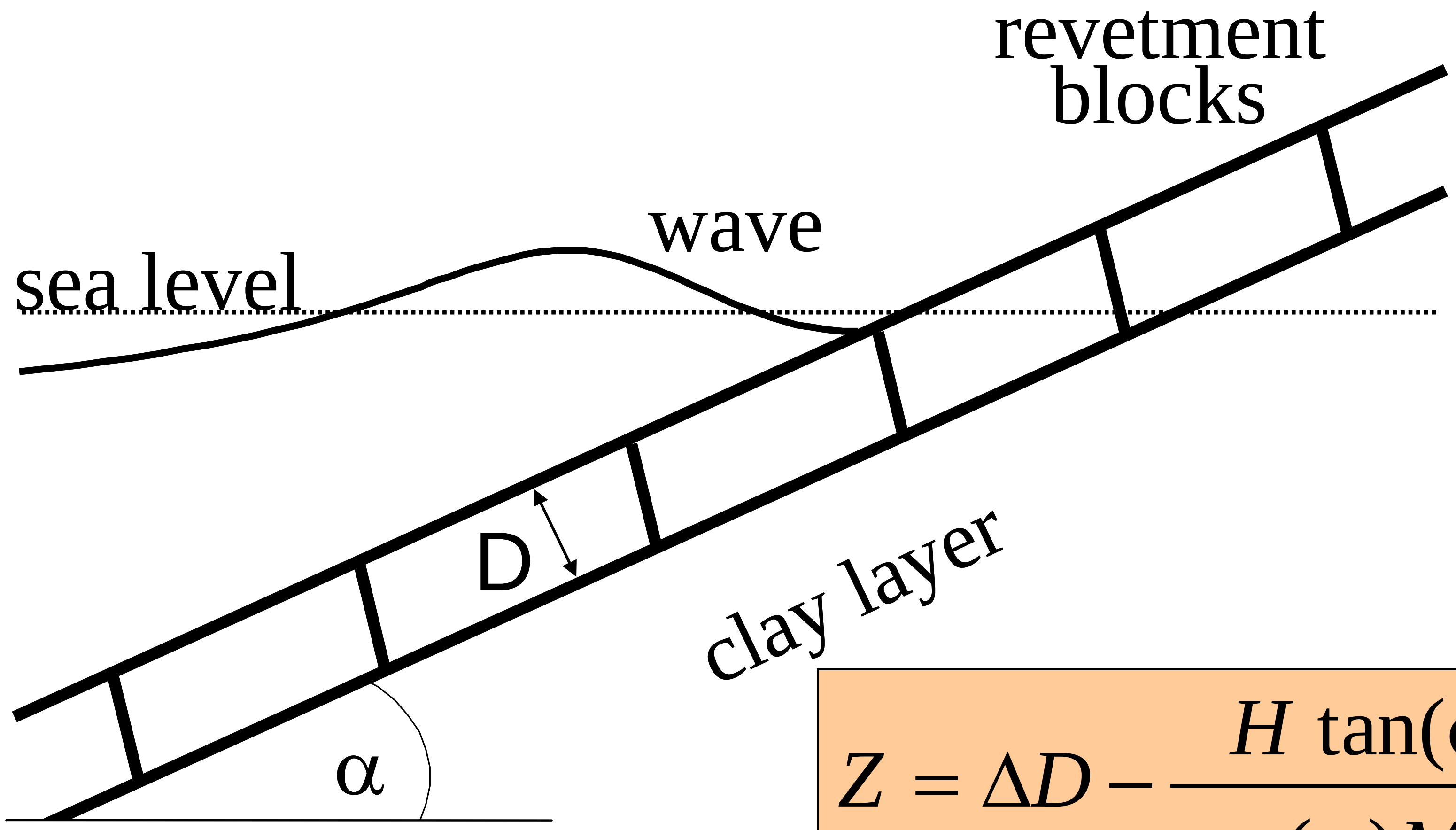
Pinching to a zero-variance interval



Pinching to other targets

- Pinchings need not be to *particular* targets, e.g., a single point or a precise distribution
- Traditionally, the results of pinching to various targets were *averaged* (called “freezing”), but that approach erases the effect of uncertainty
- Instead, we ask: what is the *range* of observed reductions in uncertainty as the pinching target is varied over many different possibilities?

Case study: dike reliability



$$Z = \Delta D - \frac{H\tan(\alpha)}{\cos(\alpha) M \sqrt{s}}$$

The inputs

Relative density of the revetment blocks
$\Delta = [1.60, 1.65]$

Block thickness
$D = [0.68, 0.72]$ meters

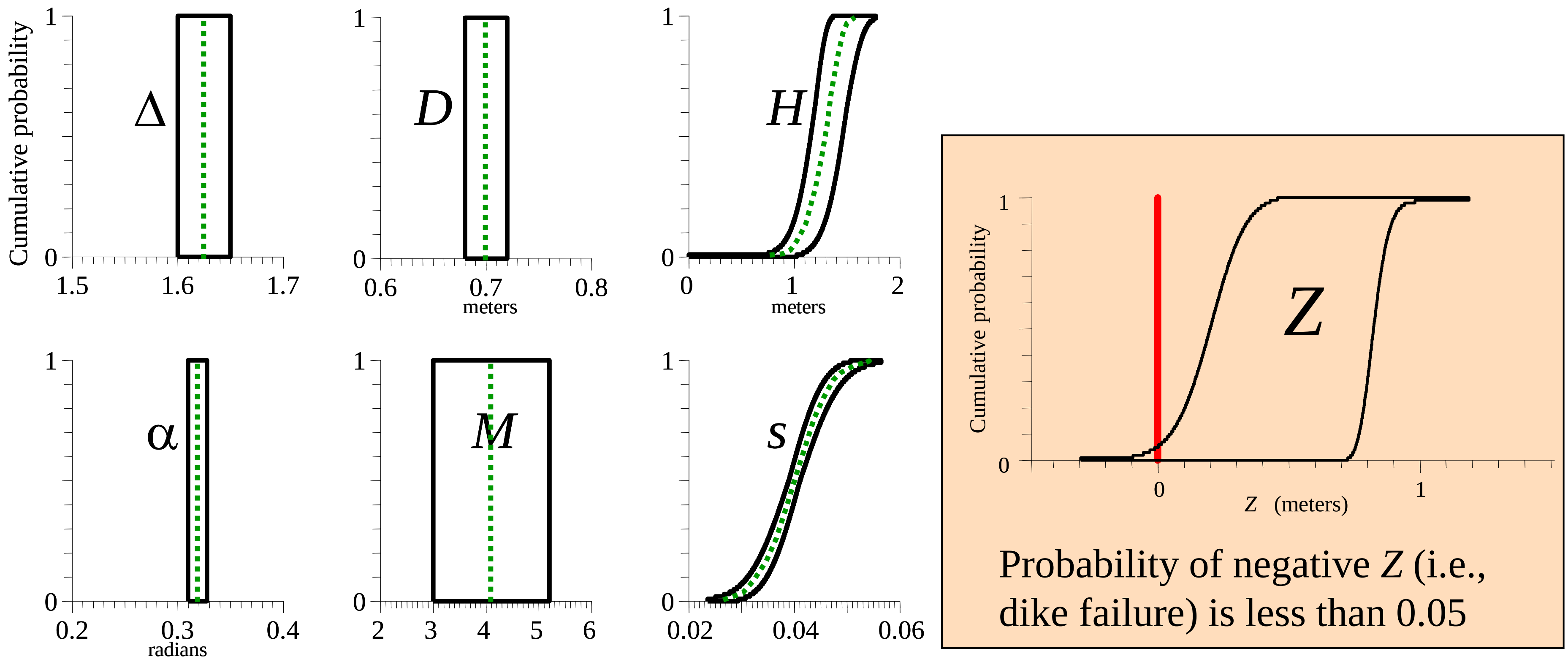
Slope of the revetment
$\alpha = \text{atan}([0.32, 0.34]) = [0.309, 0.328]$ radians

Analysts' "model uncertainty" factor
$M = [3.0, 5.2]$

Significant wave height (average of the highest third of waves)
$H = \text{weibull}([1.2, 1.5]$ meters, $[10, 12])$

Offshore peak wave steepness
$s = \text{normal}([0.039, 0.041], [0.005, 0.006])$

Analysis



Probability of negative Z (i.e., dike failure) is less than 0.05

Nominal pinching targets are dotted

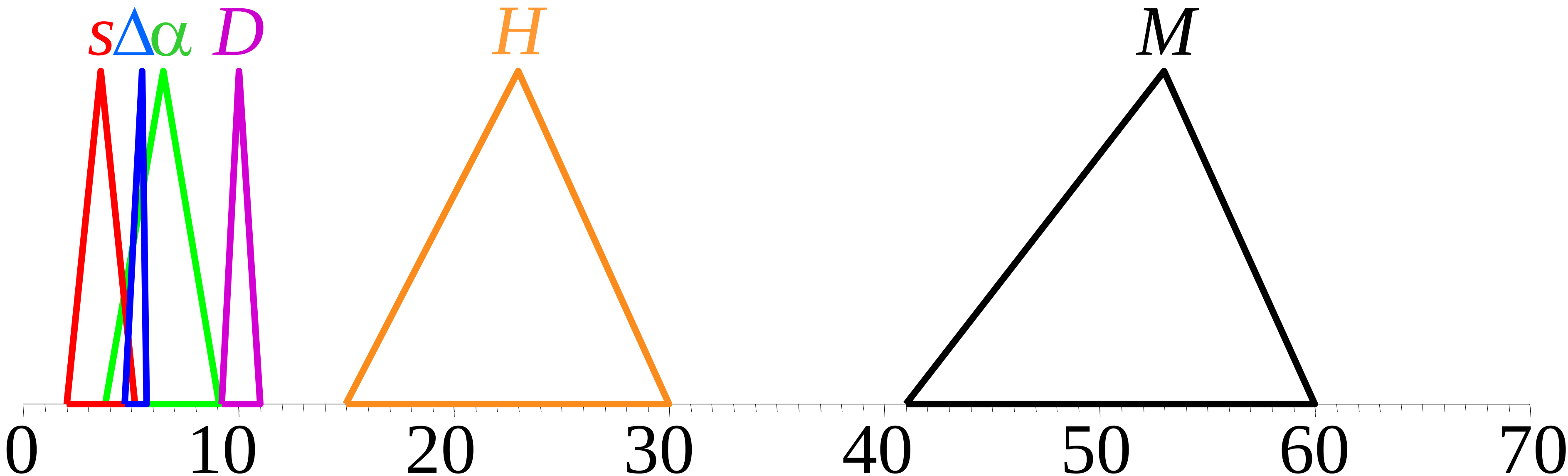
Percent reduction of uncertainty

Input	*Nominal pinching*	*All possible pinchings*
Δ	5.5	[4.7, 5.7]
D	10.0	[9.2, 11.0]
M	53.0	[41.0, 60.0]
α	6.5	[3.8, 9.1]
H	23.0	[15.0, 30.0]
s	3.6	[2.0, 5.2]

Empirical study rank order: M H D α Δ s

On the same axis

(intervals depicted as triangles)



Empirical study rank order: $M \;\; H \;\; D \;\; \alpha \;\; \Delta \;\; s$

Caveat

Omitting some variables from a sensitivity analysis ("shortlisting") because their uncertainties are small is a bad idea.

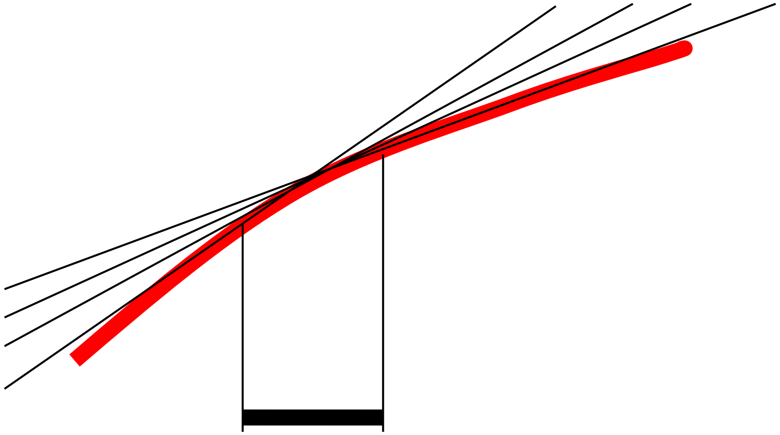
Doing this reduces dimensionality,
but it also *erases uncertainty.*

What about engineering control?

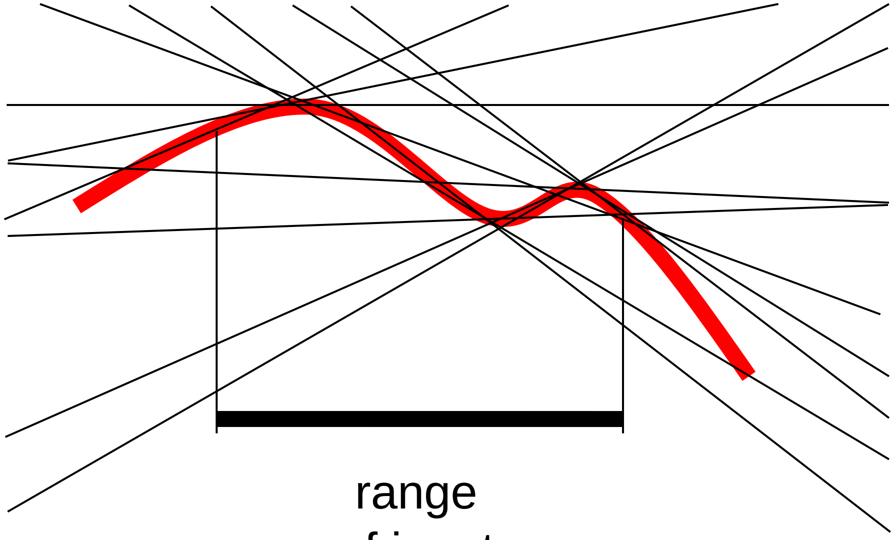
Sensitivity analysis <u>with</u> p-boxes

- Local sensitivity via derivatives
- Explored macroscopically over the uncertainty in the input
- Describes the ensemble of tangent slopes to the function over the range of uncertainty

Monotone function



Nonlinear function



Local derivatives

$$\frac{\partial Z}{\partial \Delta} = D$$

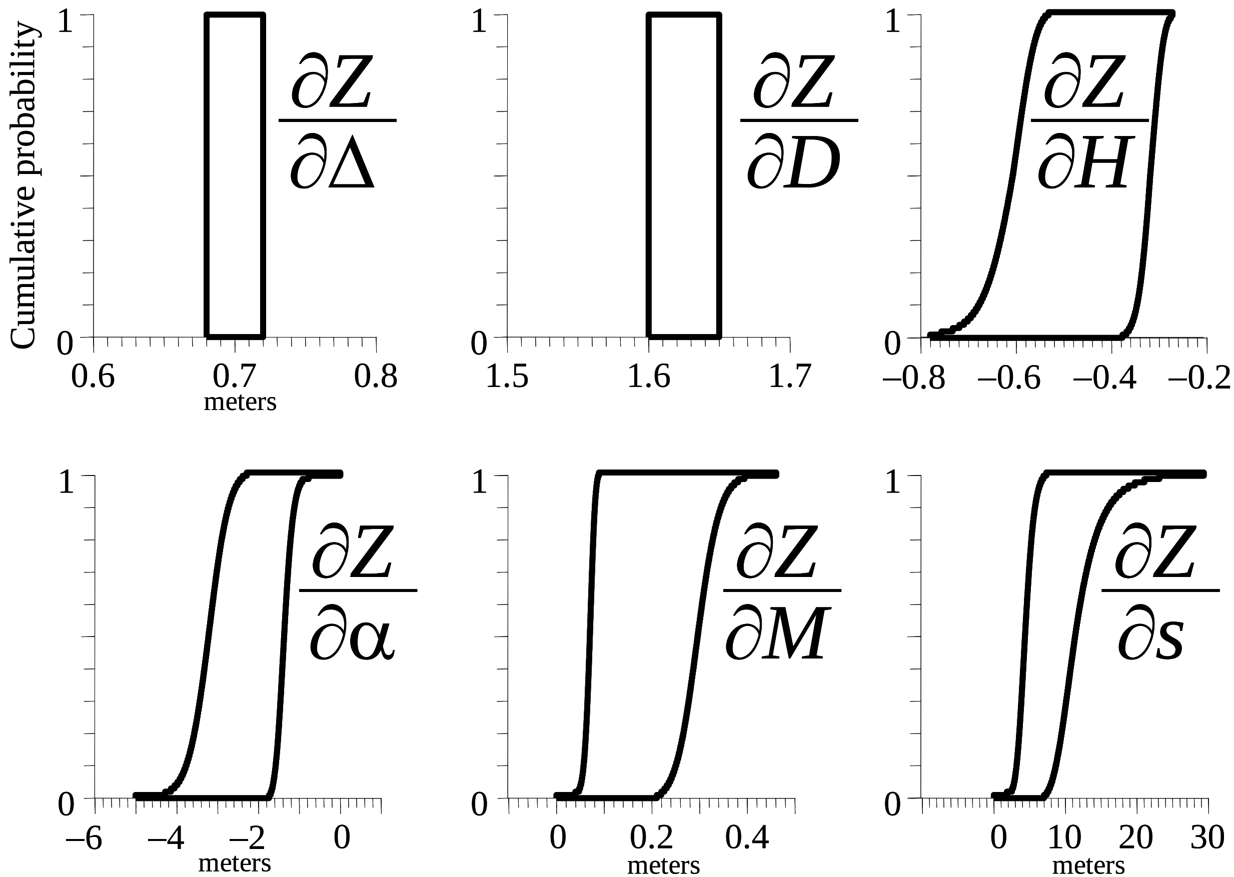
$$\frac{\partial Z}{\partial D} = \Delta$$

$$\frac{\partial Z}{\partial \alpha} = \frac{-H\left(1+\sin^2(\alpha)\right)}{\cos^3(\alpha) M \sqrt{s}}$$

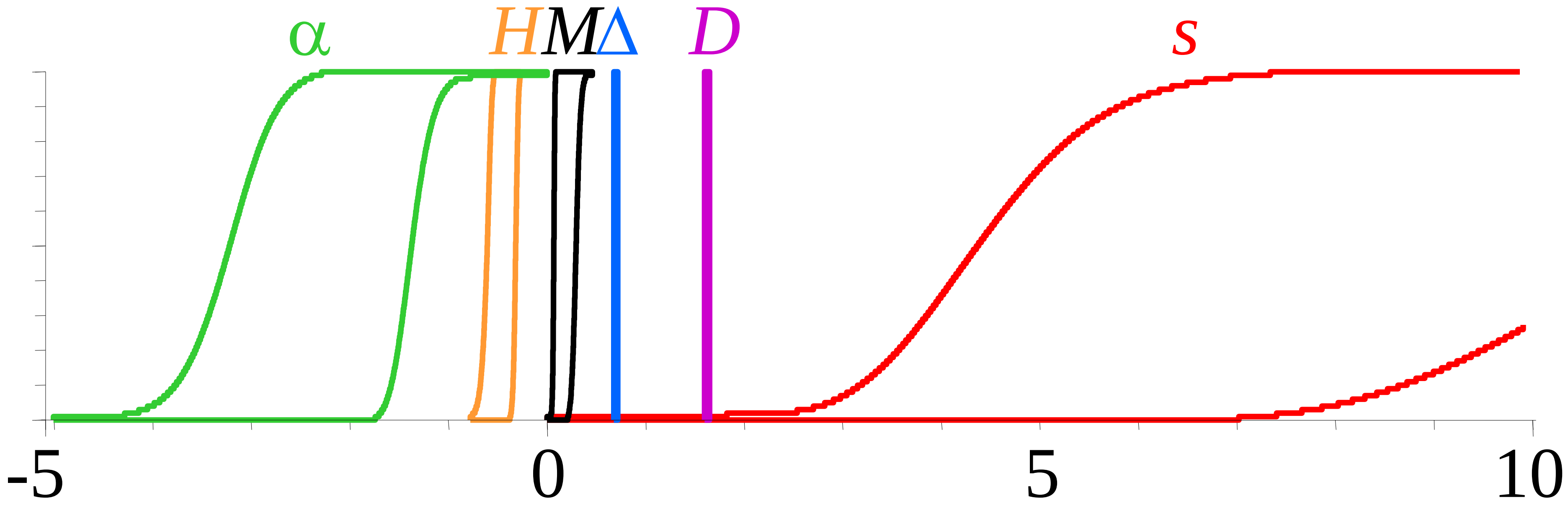
$$\frac{\partial Z}{\partial M} = \frac{H \tan(\alpha)}{\cos(\alpha) M^2 \sqrt{s}}$$

$$\frac{\partial Z}{\partial H} = \frac{-\tan(\alpha)}{\cos(\alpha) M \sqrt{s}}$$

$$\frac{\partial Z}{\partial s} = \frac{H \tan(\alpha)}{2\cos(\alpha) M s^{3/2}}$$



On the same axis



Engineering control rank order: s α D Δ H M

Completely different rankings

- The uncertainty reduction in hypothetical pinching assesses the possible effect of more or better data
- Evaluating local derivates as p-boxes tells how effective control or management can be

Best for study

Empirical study: $M \;\; H \;\; D \;\; \alpha \;\; \Delta \;\; s$

Engineering control: $s \;\; \alpha \;\; D \;\; \Delta \;\; H \;\; M$

Best for control

Different ‘sensitivity’ questions

✓ Empirical planning
  – What variables need study to reduce uncertainty?

✓ Engineering control
  – What variables can be modified to change the result?

✓ Robustness analysis
  – How robust are the results of the assessment?

✘ Tracking analysis
  – Which input combinations yield extreme results?

Backcalculation

Backcalculation

- Cleanup and remediation planning requires backcalculation

- How can we untangle the expression

$$A+B = C$$

when we know A and C, and need B?

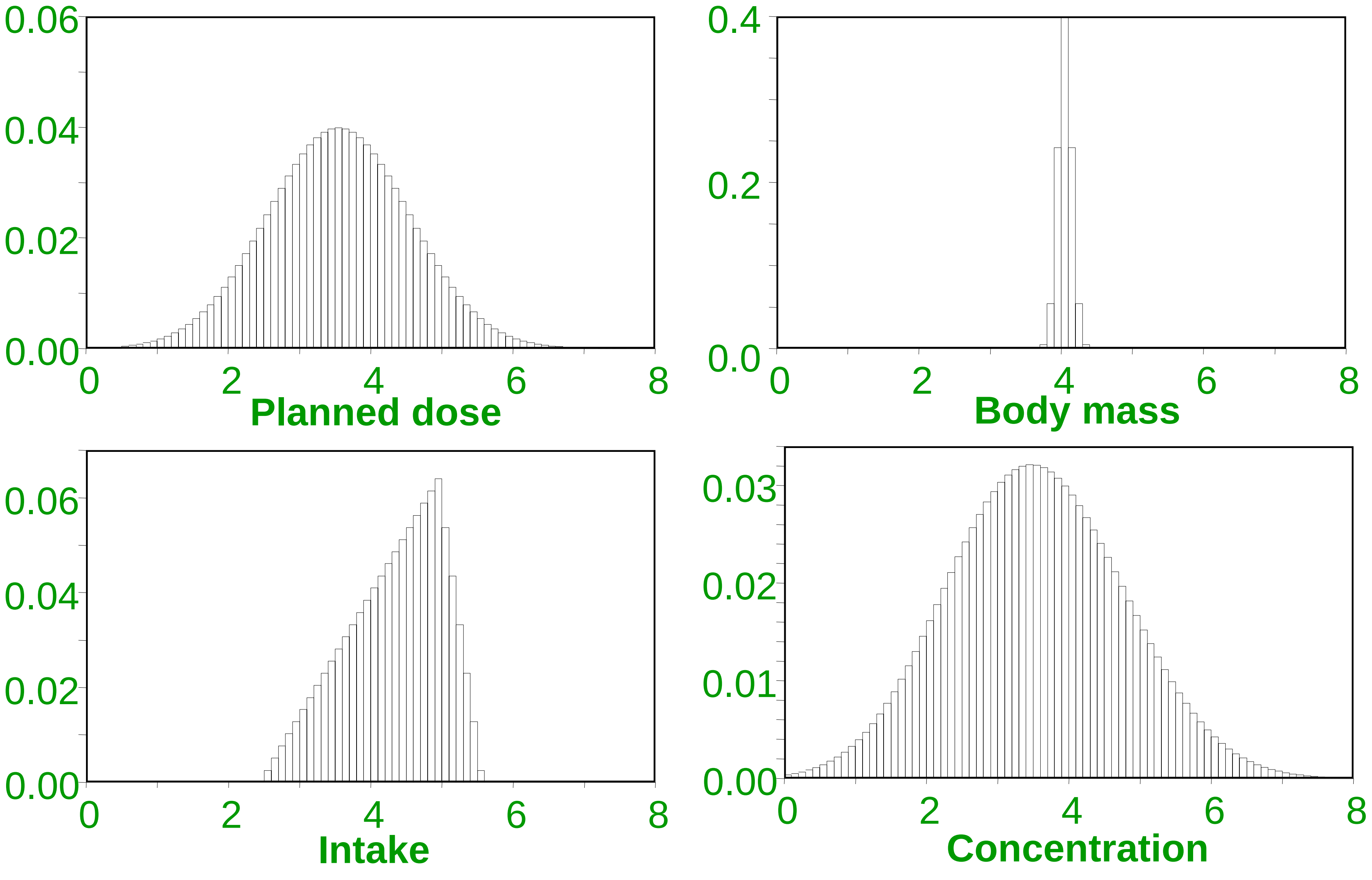
Can't just invert the equation

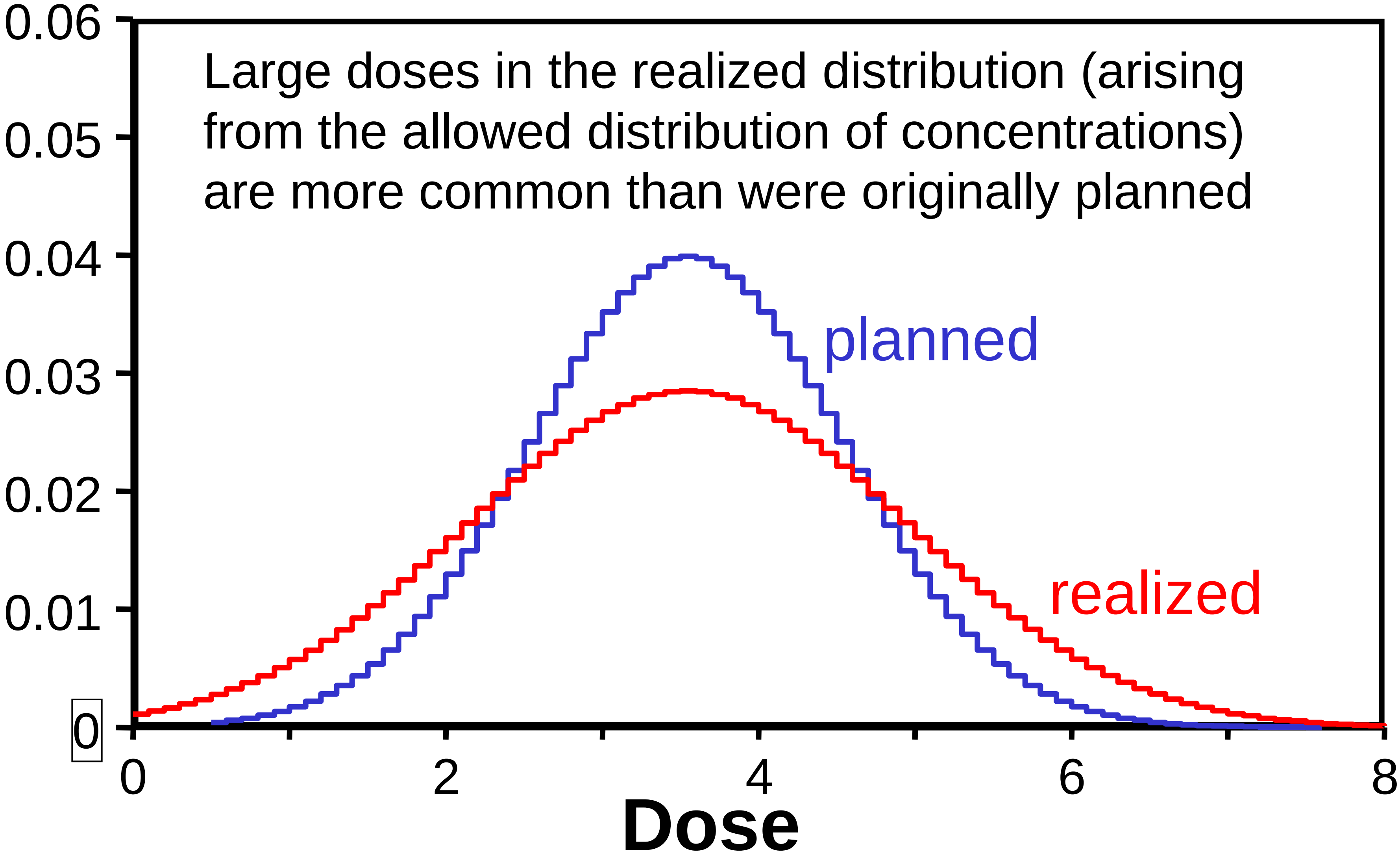
prescribed unknown known

$$Dose = Concentration \times Intake \,/\, Bodymass$$

$$Concentration = Dose \times Bodymass \,/\, Intake$$

When concentration is put back into the forward equation, the resulting dose is wider than planned


0.06
0.04
0.02
0.00
0
2
4
6
8
Planned dose
0.4
0.2
0.0
0
2
4
6
8
Body mass
0.06
0.04
0.02
0.00
0
2
4
6
8
Intake
0.03
0.02
0.01
0.00
0
2
4
6
8
Concentration



Large doses in the realized distribution (arising
from the allowed distribution of concentrations)
are more common than were originally planned
planned
realized
0.06
0.05
0.04
0.03
0.02
0.01
0
0
2
4
6
8
Dose

Normal approximation

- If $A+B=C$, compute B as $C–A$ under the assumption the correlation between A and C is r = sd(A)/sd(C)

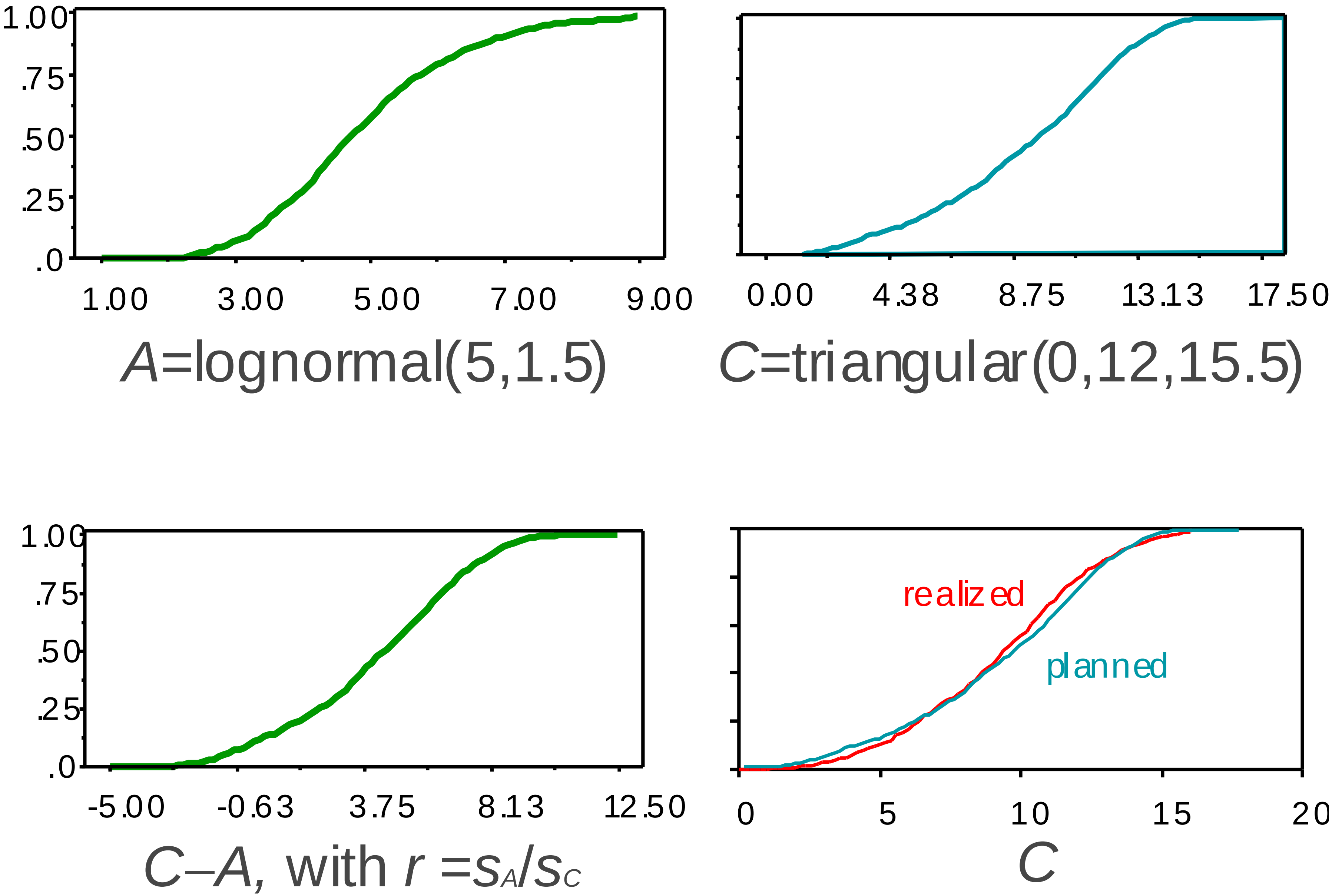
To simulate normal deviates with correlation r, compute

$$Y_1 = Z_1 \times \sigma_1 + \mu_1$$
$$Y_2 = (rZ_1 + Z_2 \sqrt{(1-r^2)}) \times \sigma_2 + \mu_2$$

where Z_1 and Z_2 are independent standard normal deviates, and μ and σ denote the respective desired means and standard deviations

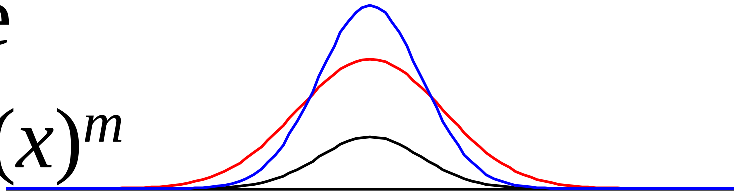
- Uses Pearson correlation (not rank correlation)
- Good for multivariate normal, and maybe more

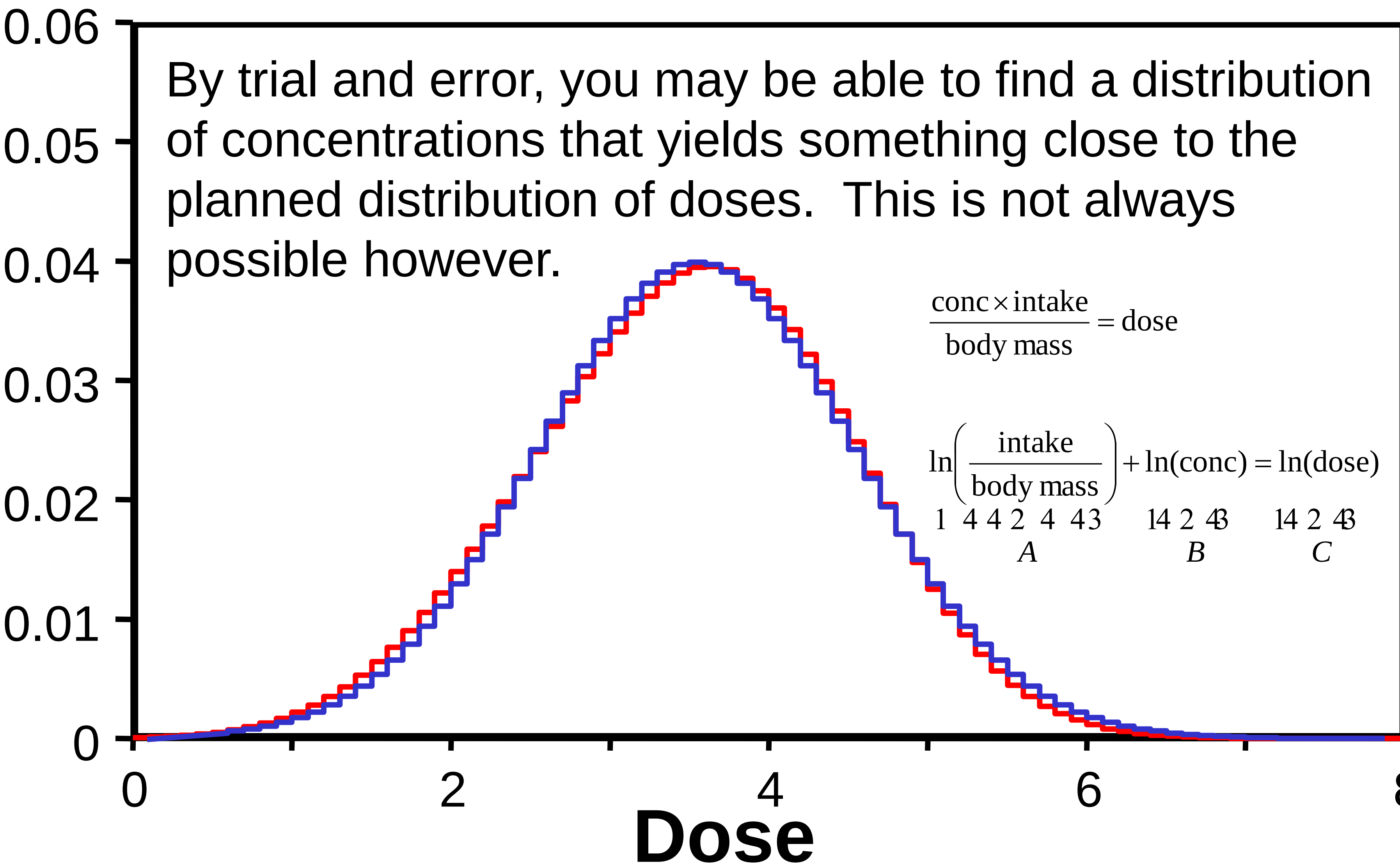
Normal approximation with non-normal distributions



Iterative (trial & error) approach

- Initialize B with $C–A$
- This distribution is too wide
- Transform density $p(x)$ to $p(x)^m$
- Rescale so that area remains one
- Whenever $m > 1$ dispersion decreases
- Vary m until you get an acceptable fit



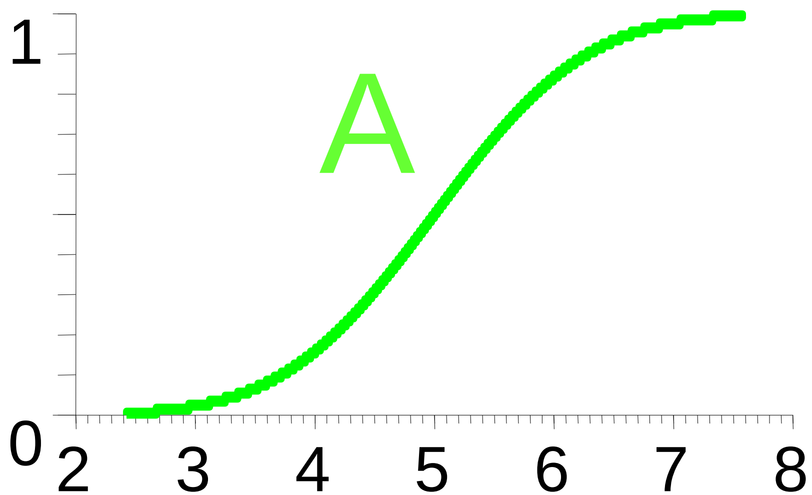

By trial and error, you may be able to find a distribution of concentrations that yields something close to the planned distribution of doses. This is not always possible however.
$$\frac{\text{conc} \times \text{intake}}{\text{body mass}} = \text{dose}$$
$$\underbrace{\ln\left(\frac{\text{intake}}{\text{body mass}}\right)}_{A} + \underbrace{\ln(\text{conc})}_{B} = \underbrace{\ln(\text{dose})}_{C}$$
0.06
0.05
0.04
0.03
0.02
0.01
0
0
2
4
6
8
Dose

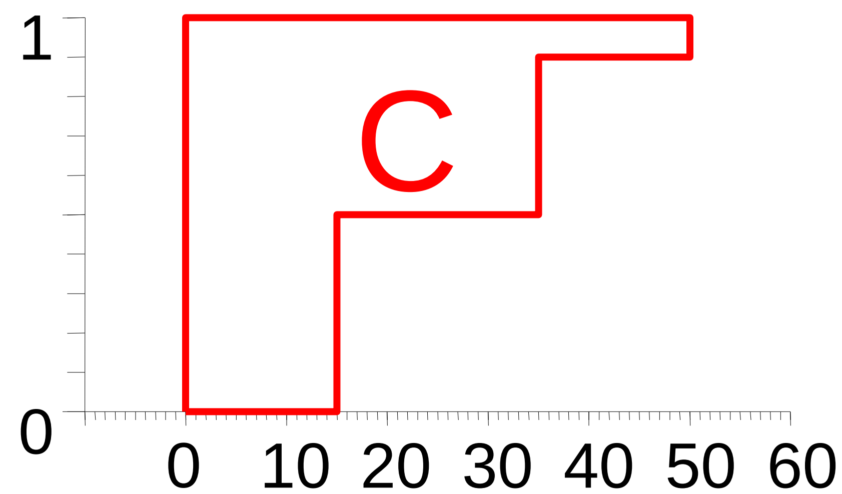
Backcalculation with p-boxes

Suppose $A + B = C$, where

A = normal(5, 1)

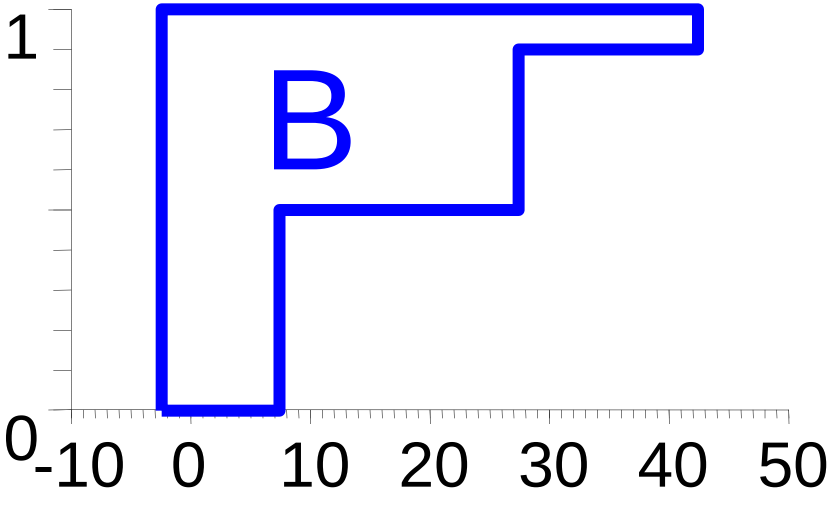
$C = \{0 \leq C$, median ≤ 1.5, 90th %ile ≤ 35, max $\leq 50\}$





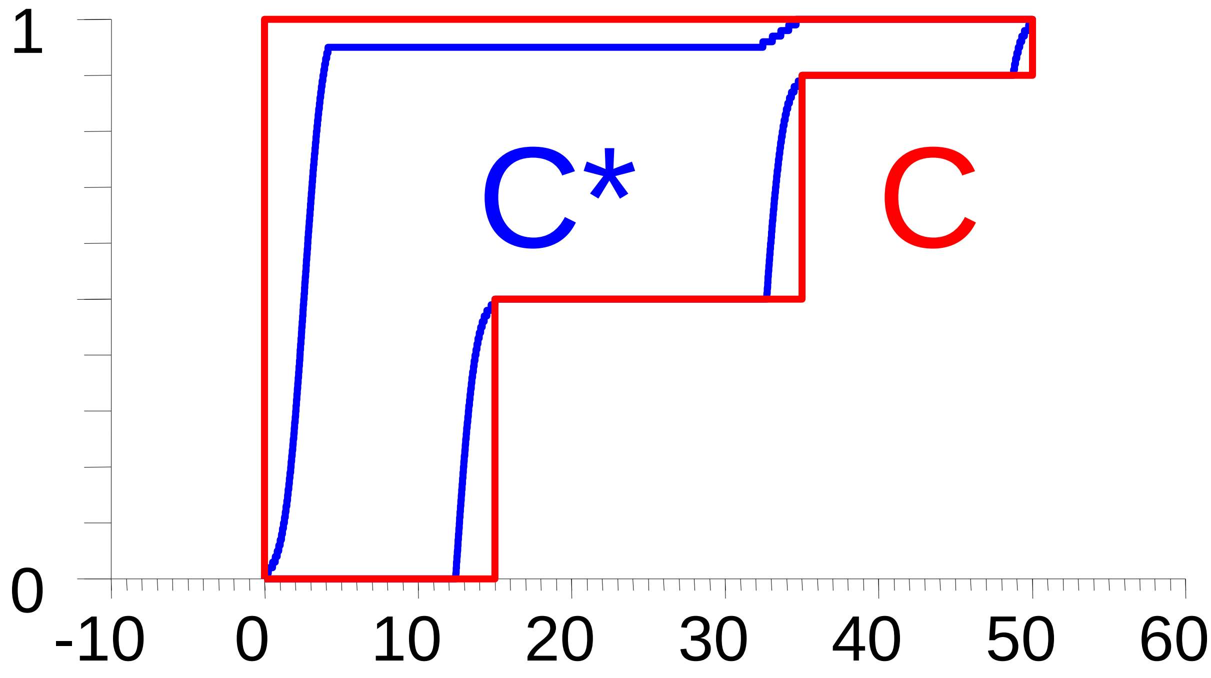
Getting the answer

- The backcalculation algorithm basically reverses the forward convolution
- Any distribution totally inside *B* is sure to satisfy the constraint … it's a "kernel"



Check it by plugging it back in

$\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{C}^* \subseteq \boldsymbol{C}$



Precise distributions don’t work

- Precise distributions can’t express the target
- A concentration distribution giving a prescribed distribution of doses seems to say we *want* some doses to be high
- Any distribution to the left would be better
- A p-box on the dose target expresses this idea

Backcalculation algebra

- Can define untanglings for all basic operations

  e.g., if $A \times B = C$, then $B = \exp(\text{backcalc}(\ln A, \ln C))$

- Can chain them together for big problems
- Assuming independence *widens* the result
- Repeated terms need special strategies

Conclusion

- Planning cleanup requires backcalculation
- Monte Carlo methods don’t generally work except in a trial-and-error approach
- Can express the dose target as a p-box

Imprecise statistics

Small sample size

- Student's *t* statistic introduced in 1908
- Statistics has spent a century developing analyses in which sample size is limiting
- But sample sizes are not so small anymore
  - Financial data
  - Continuous mechanized sampling
  - Satellite imagery and other mass collections
  - Commercial data
  - Social media
  - Internet of Things (30 trillion sensors feeding the web)

Other uncertainties

- Sample size will always be an issue
- But it may not always be the *only* issue anymore
- Other issues become important as sample sizes grow
  - Measurement imprecision (mensurational uncertainty)
  - Missingness and censoring
  - Model uncertainty, non-stationarity, etc.
- Many believe it's always better to collect more samples than to improve the precision of samples
  - This is *not* true
  - Believing this creates suboptimal experimental designs

Measurements aren't reals

- Real line is a poor model for measurements
  - "Measure" by comparing readings against a scale
  - Almost all real values cannot be measurements
- All real measurements have uncertainties
  - A real value cannot express this
- Real values have infinitely many zeros after the last decimal place
  - Such precision is never achievable in the real world
- The real line is totally ordered
  - But a measurement may not be only larger, smaller or the same as another

Missing data

- Traditional methods assume MCAR or MAR
- These assumptions are not always reasonable
- Assuming them anyway leads to wrong answers
- Correctly accounting for missingness can yield *dilation*, in which your uncertainty increases even when you increase sample size
  - If a temperature sensor fails from extremes in either direction, a missing value may mean the temperature is much higher or much lower than you thought

Censoring

- Traditional methods are decidedly bad and can be grossly misleading
- Likelihood strategies make assumptions that may not be tenable
  - Can produce unreasonable results

Interval data

- Calculating variances, *t*-statistics, etc. for data sets that contain intervals are NP-hard problems
- But various special cases are quite easy
  - Censored
  - Binned
  - Same precision
  - Nested
- These let us compute variance in $O(n)$ time
  - Can compute variance for 500,000 samples in less than 0.5 second on a laptop

Interval uncertainty

Incertitude

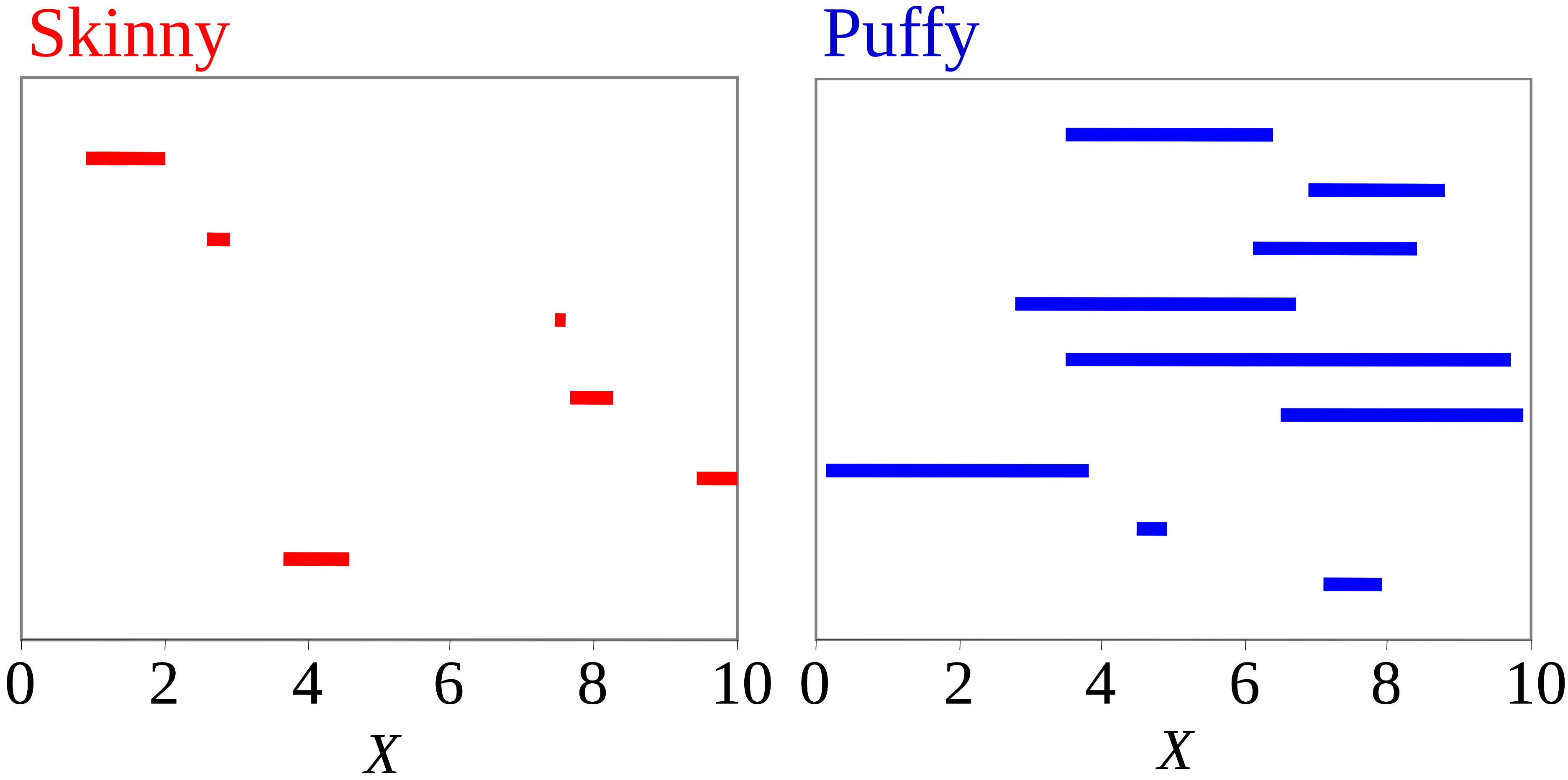
- Intermittent observations
- Plus-minus intervals
- Non-detects and data censoring
- Missing values
- Blurring for privacy or security reasons
- Bounding studies

Two approaches

- Model each interval as uniform distribution
  - Presumes different values are equally likely
  - Laplace's principle of insufficient reason
  - Calculations relatively easy, but interpretation subtle
- Model each interval as a *set* of possible values
  - Specifies no single distribution within the range
  - Theory of imprecise probabilities
  - Calculations often NP-hard, but interpretation easy

A tale of two data sets

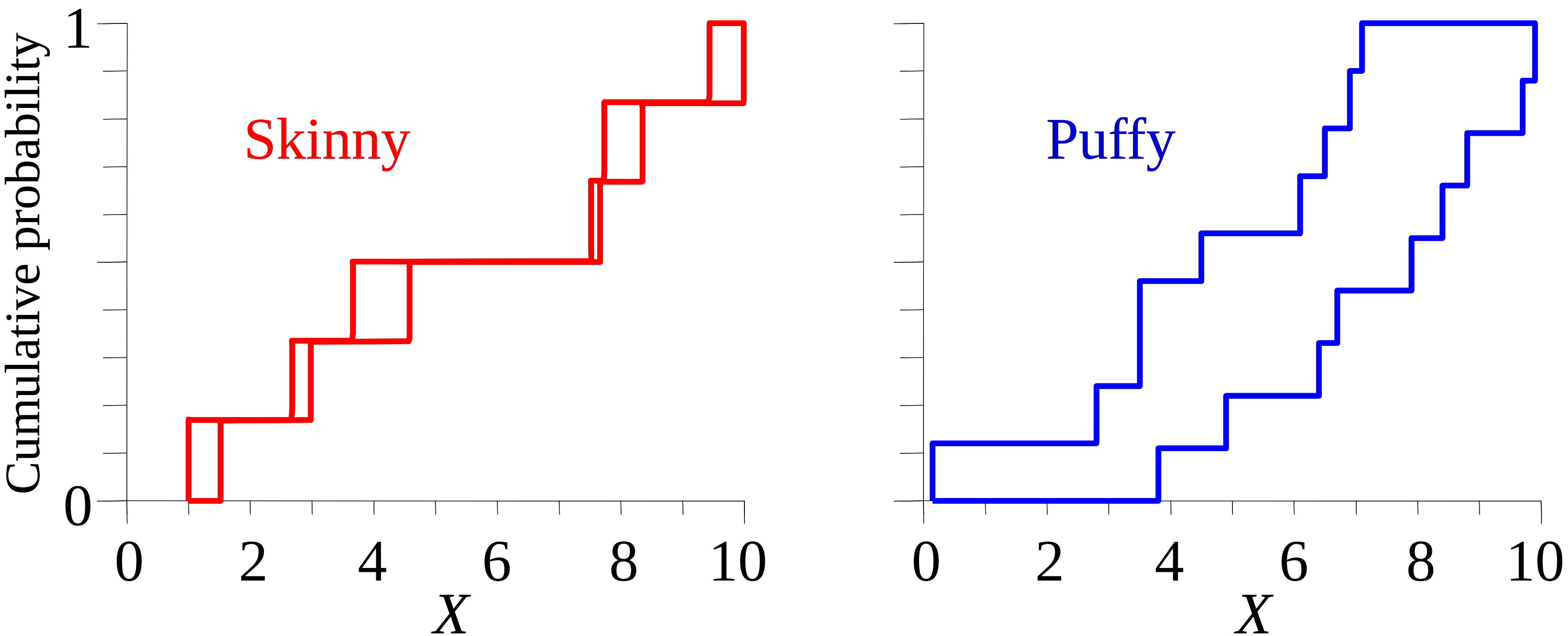
Skinny data	**Puffy data**
[1.00, 2.00]	[3.5, 6.4]
[2.68, 2.98]	[6.9, 8.8]
[7.52, 7.67]	[6.1, 8.4]
[7.73, 8.35]	[2.8, 6.7]
[9.44, 9.99]	[3.5, 9.7]
[3.66, 4.58]	[6.5, 9.9]
	[0.15, 3.8]
	[4.5, 4.9]
	[7.1, 7.9]


Skinny
Puffy
0
2
4
6
8
10
X

Empirical distribution

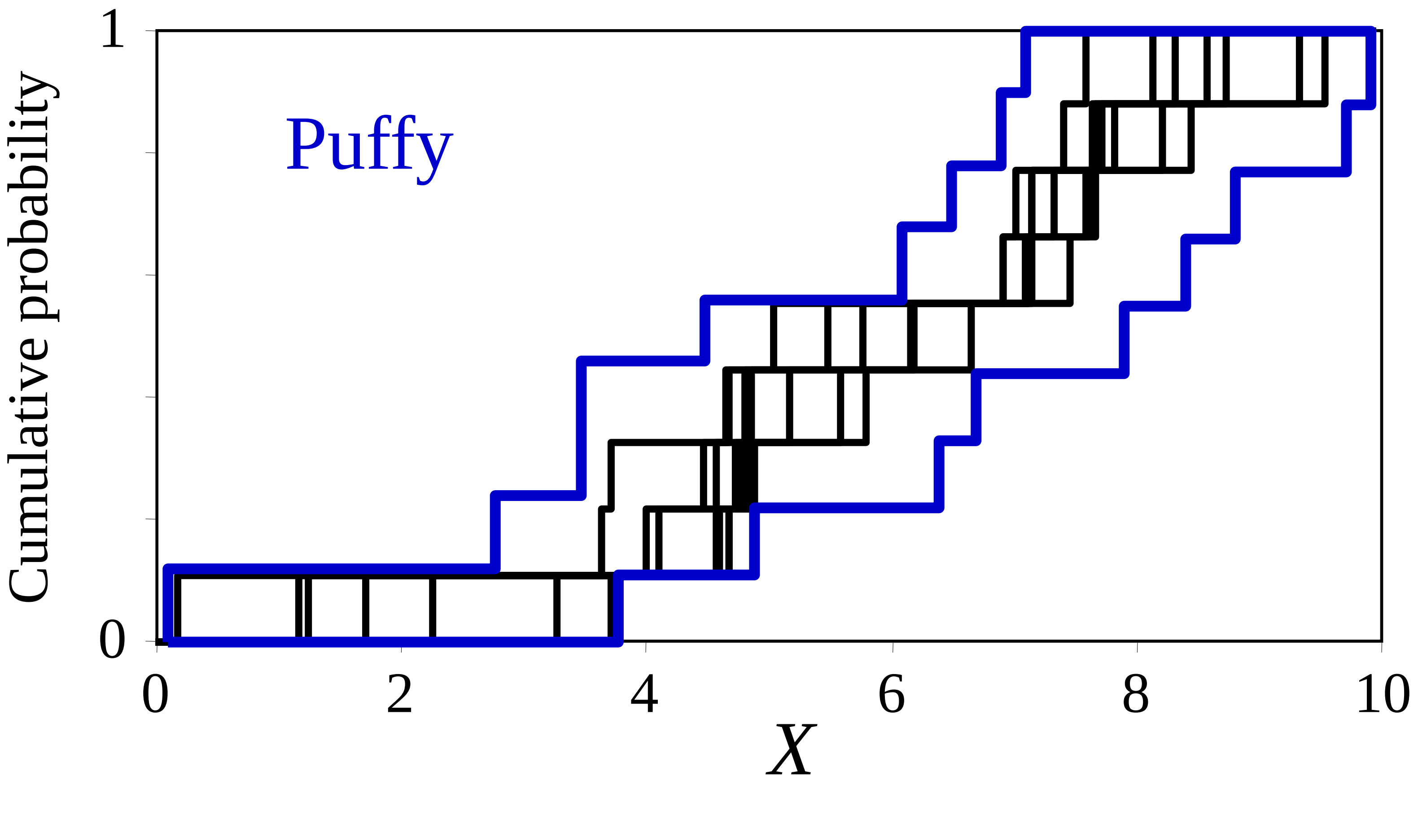
- Summary of the data themselves
- No distributional assumptions
- Uniforms approach yields a single distribution
- Intervals approach yields a probability box (i.e., a class of distributions)

Intervals approach

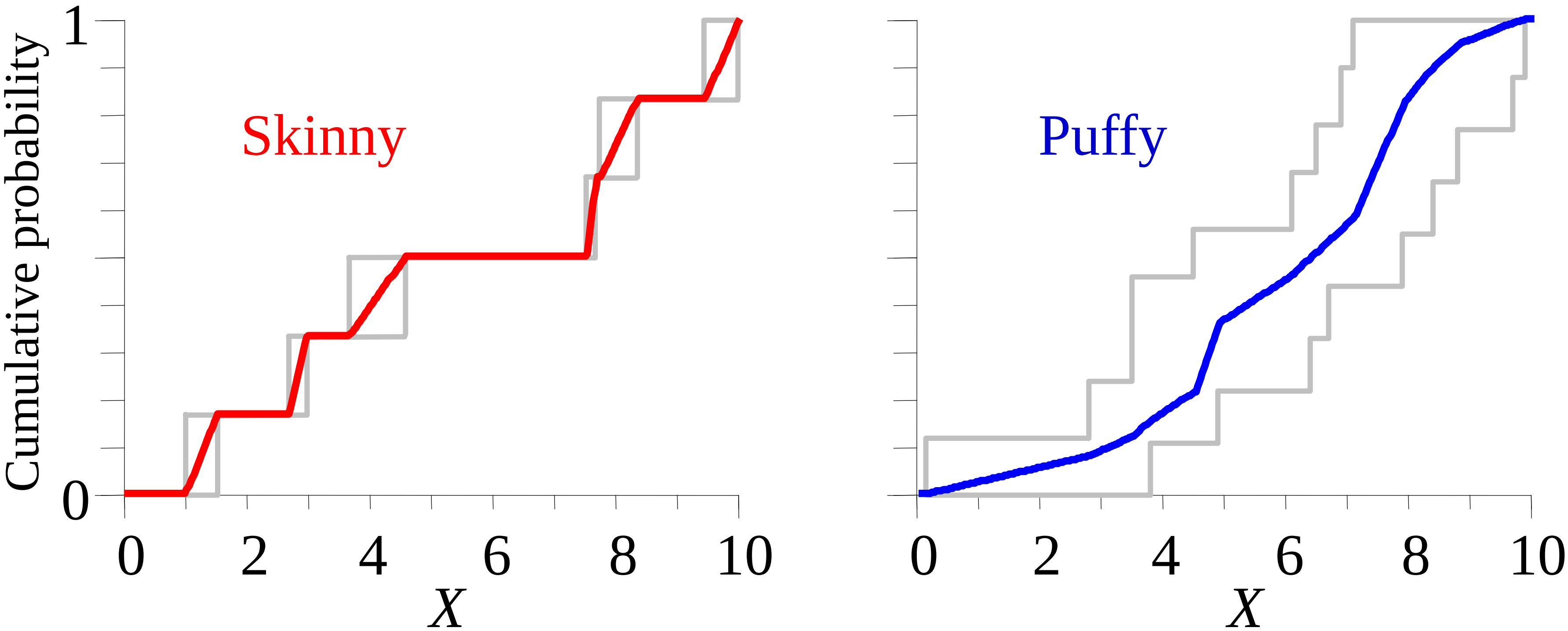


- Each side is cumulation of respective endpoints
- Represents both incertitude and variability

Uncertainty about the EDF



Uniforms approach



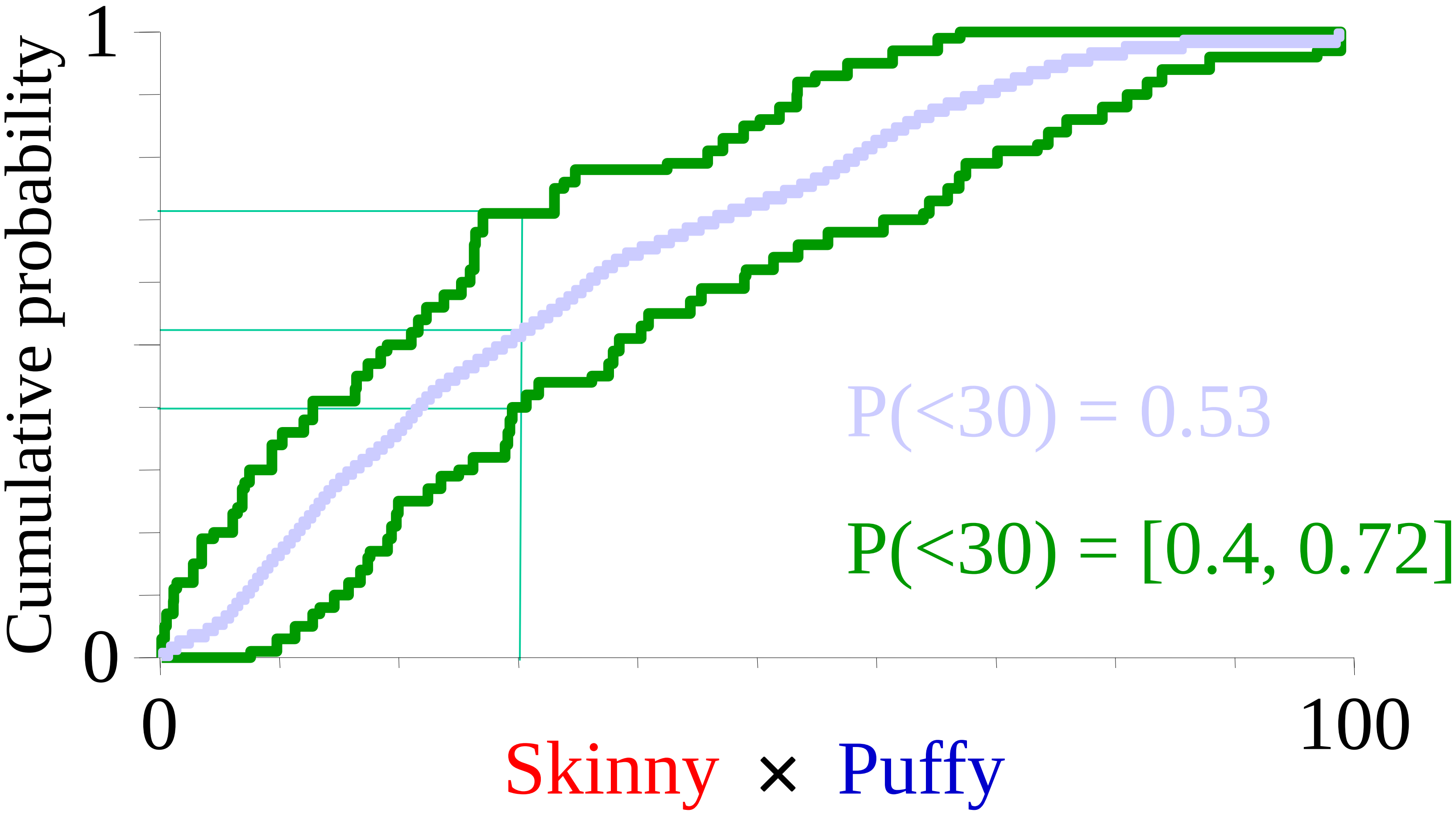
- Mixture (vertical average) of uniforms
- Conflates incertitude and variability

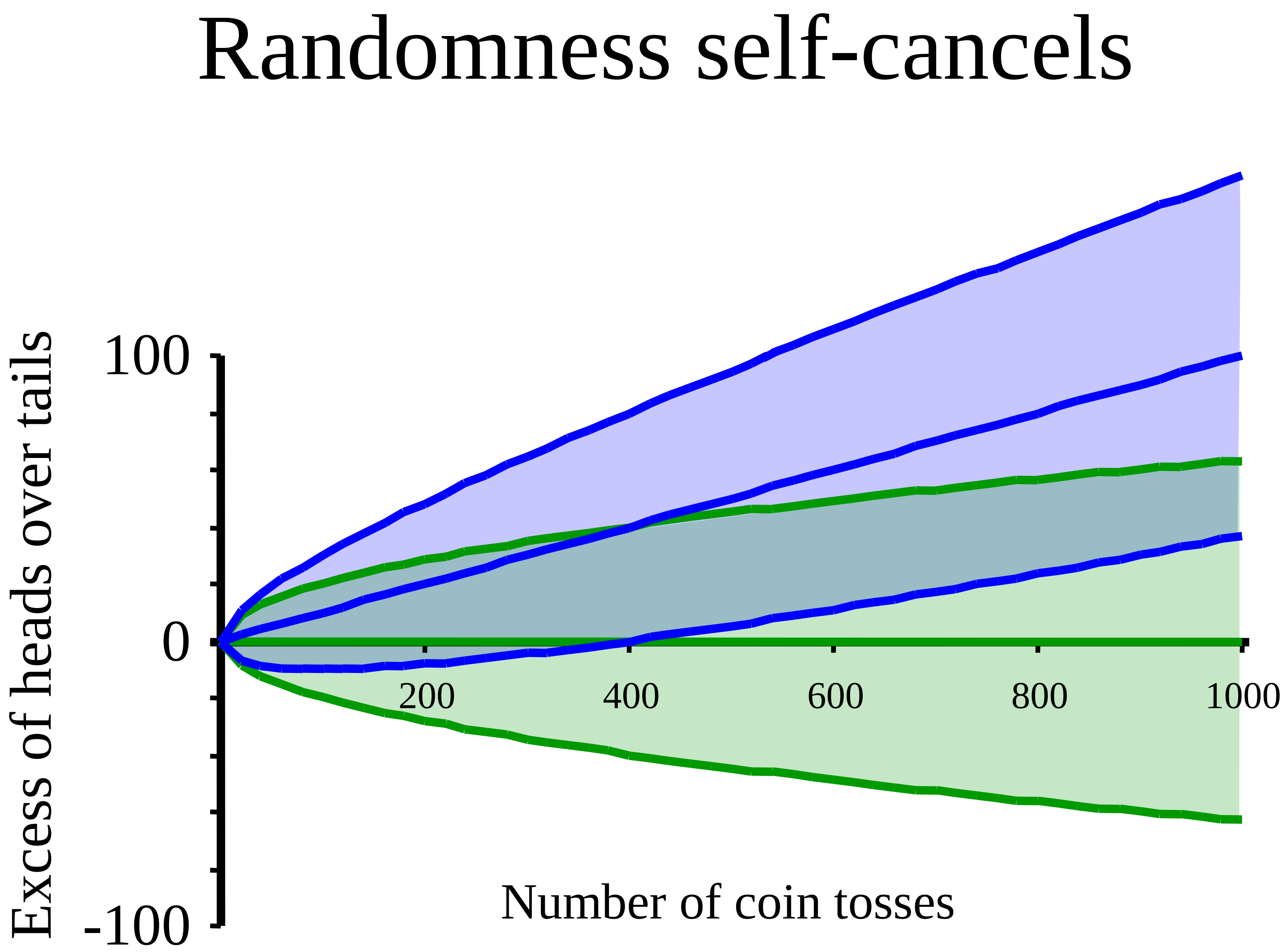
What’s the difference?

- Might be advisable to propagate the two kinds of uncertainty separately and differently

- *Example:* suppose we’re interested in the product of Skinny and Puffy…

Uniforms versus intervals approach




Randomness self-cancels
Excess of heads over tails
100
0
-100
200
400
600
800
1000
Number of coin tosses

Fitted distribution

- Assumes some *shape* for the distributions
- Uniforms approach yields a single distribution
- Intervals approach yields a probability box

Interval approach

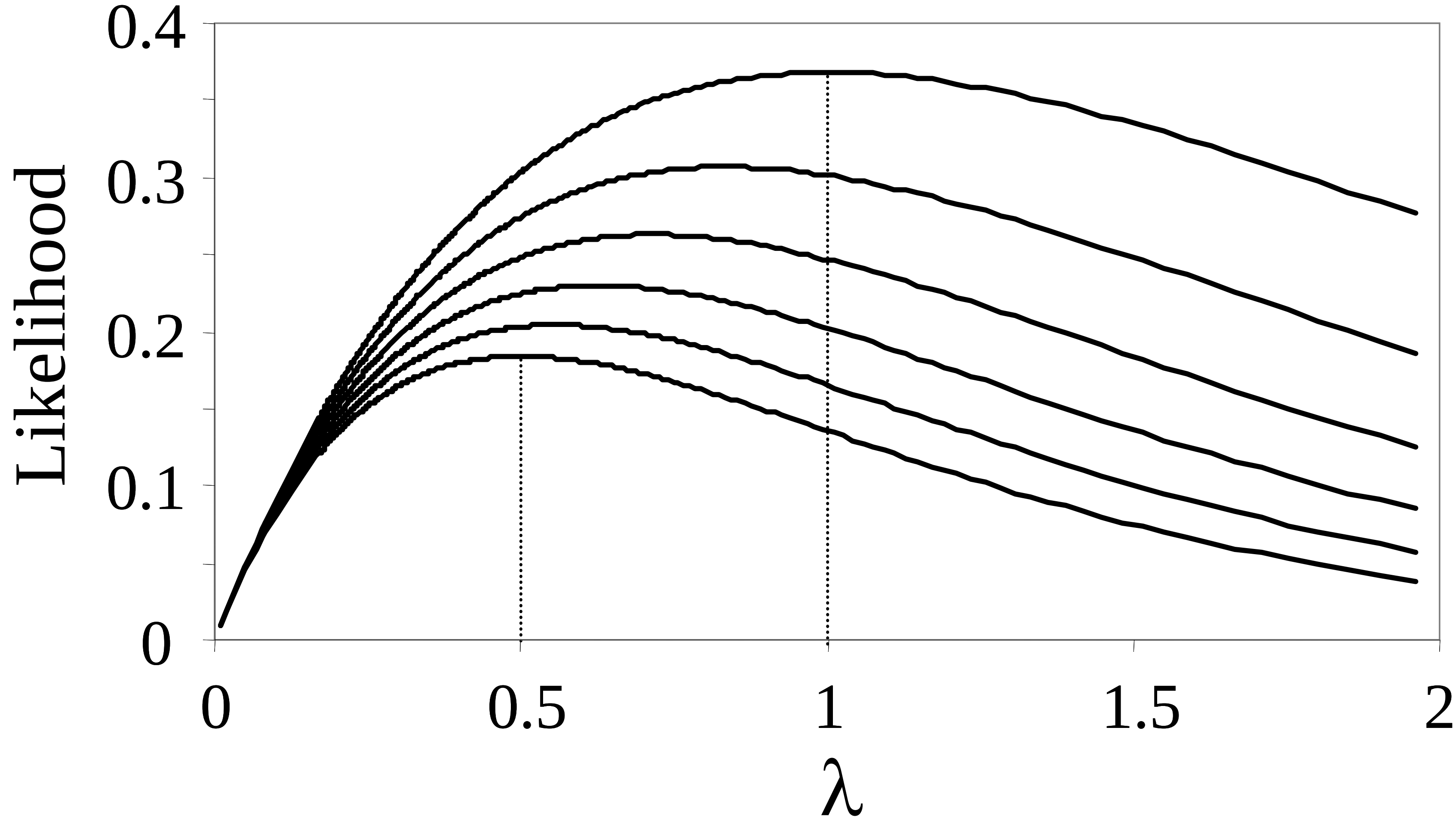
- Creates a class of maximum likelihood solutions

- Every one solves a ML problem for a set of scalar values within the respective intervals

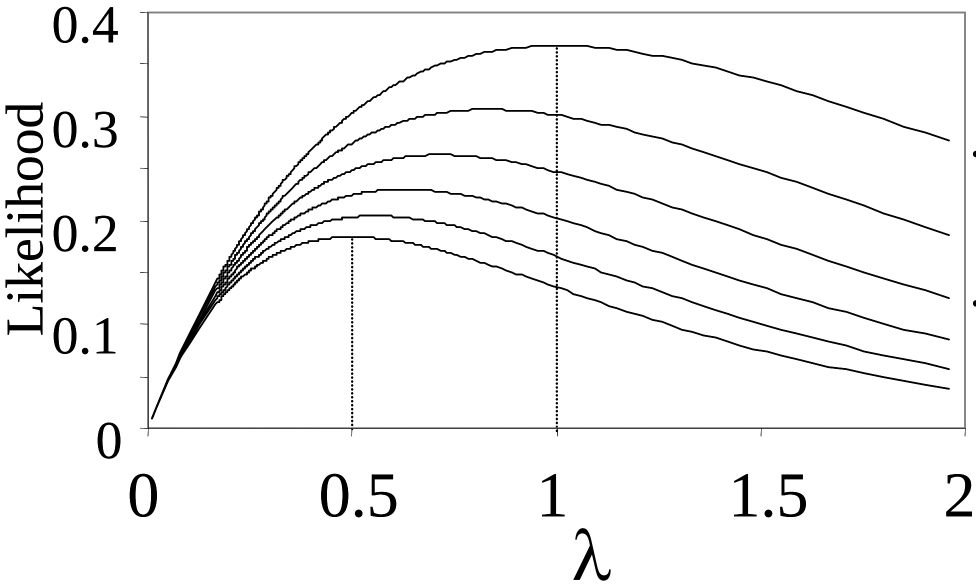
- *Example:* let's fit exponential distributions

$$L(x;\lambda) = \lambda \exp(-\lambda x) \qquad x \in [1,2]$$

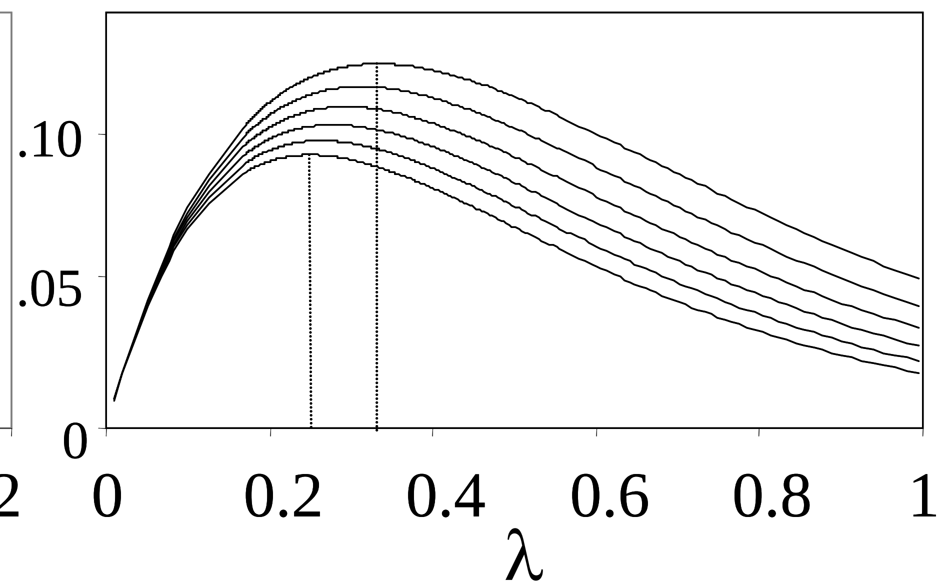
densities

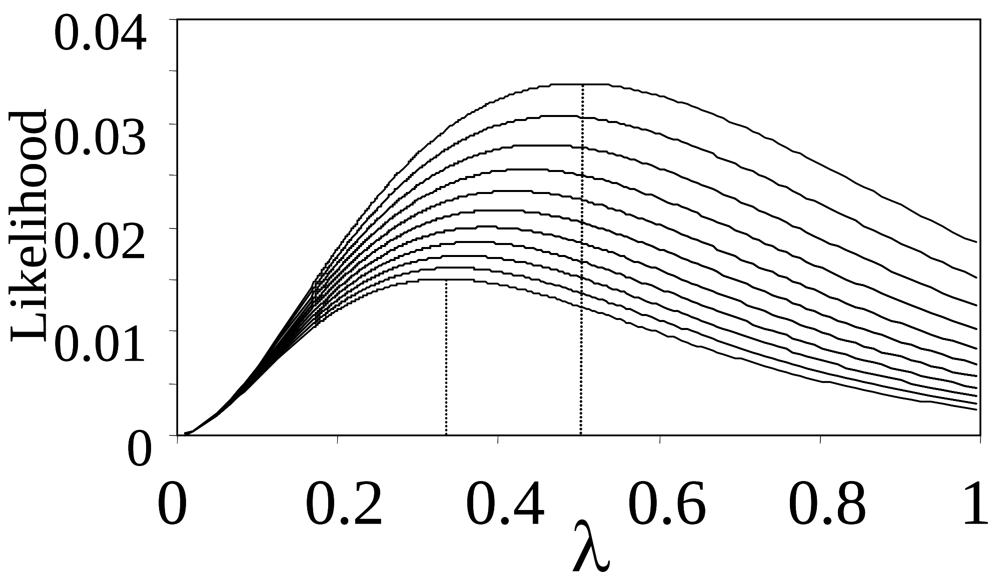


$x_1 \in [1,2]$



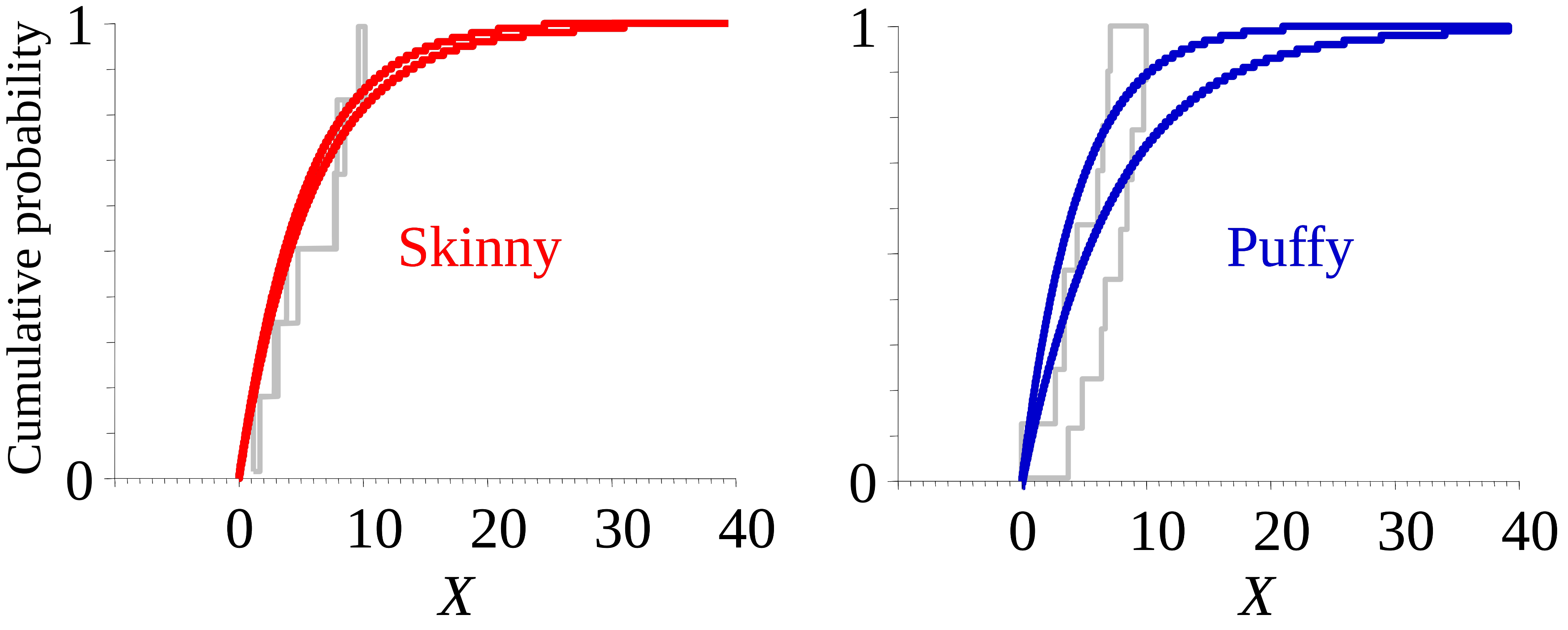
$x_2 \in [3,4]$





Pairwise multiplied if data are independent

Results from intervals approach



Maximum likelihood for censored data

Given a datum $\boldsymbol{x} = [\underline{x}, \overline{x}]$,

$$
\begin{aligned}
L(\boldsymbol{x}) &= \Pr(\underline{x} \le X \le \overline{x}) \\
&= \Pr(\overline{x} \le X) - \Pr(\underline{x} \le X) \\
&= F(\overline{x};\, \lambda) - F(\underline{x};\, \lambda)
\end{aligned}
$$

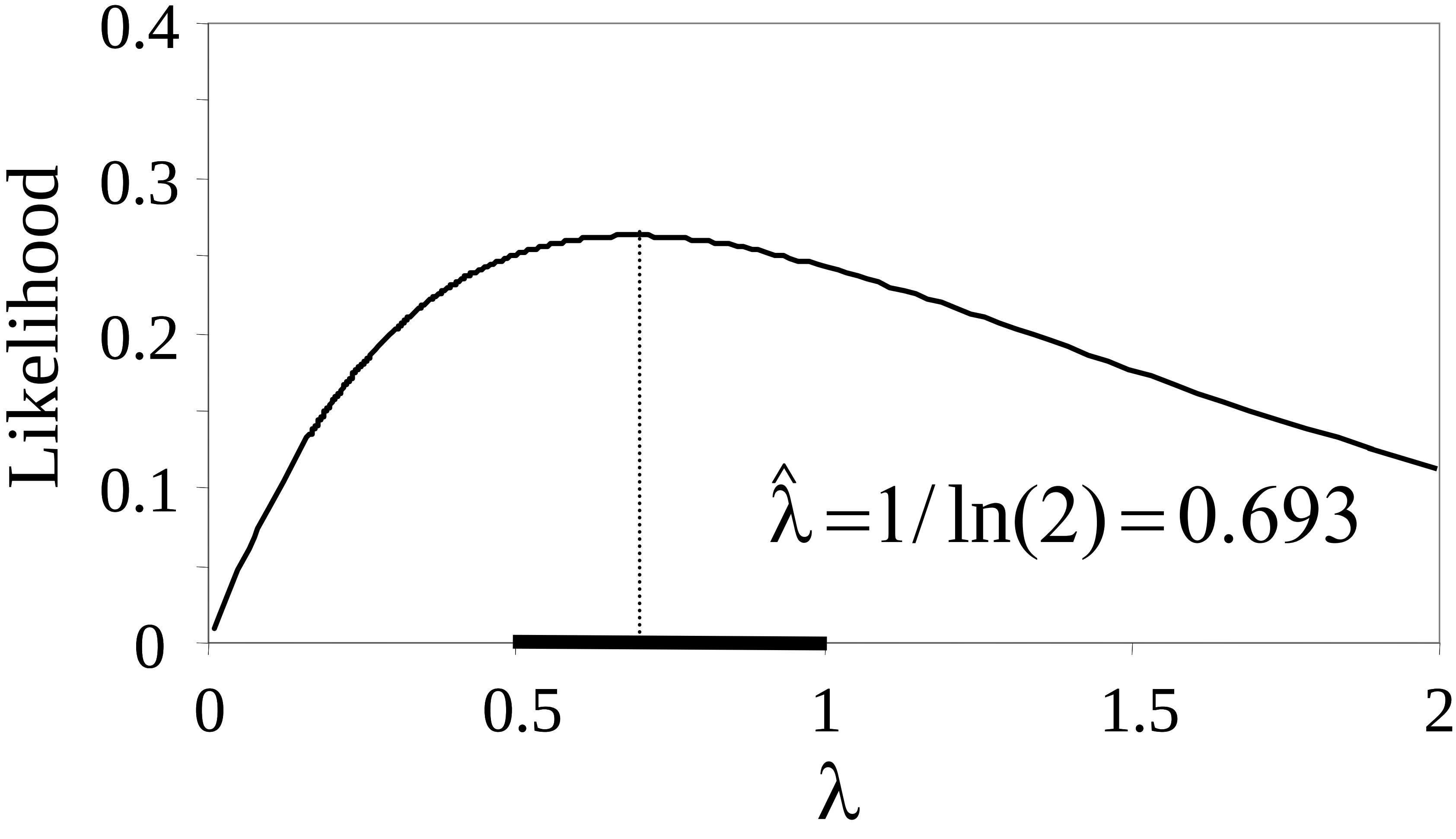
where, for an exponential distribution,

$$F(x;\lambda) = 1 - \exp\left(-\lambda x\right)$$

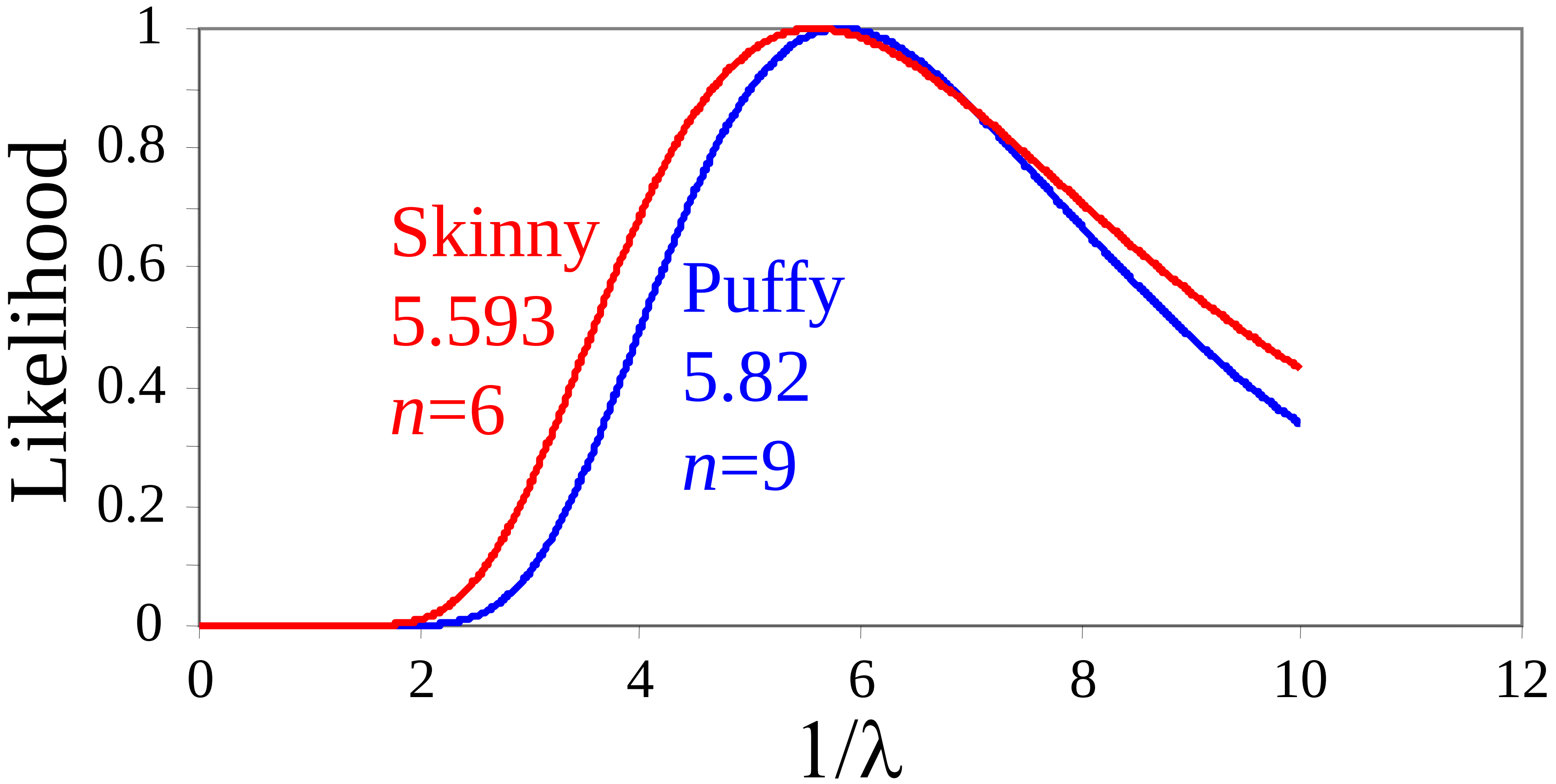
and $1/\lambda$ is the mean

Single datum $x \in [1,2]$

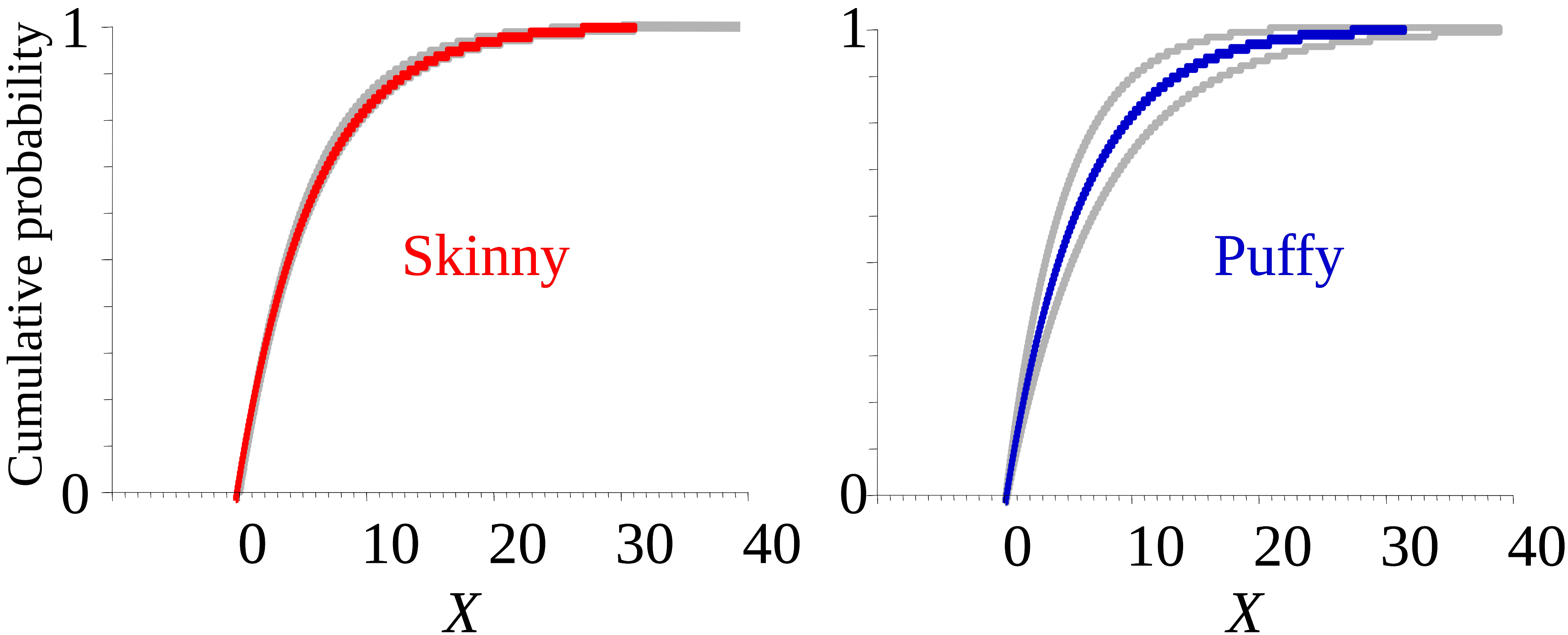
$$L(\lambda) = \exp(-\lambda) - \exp(-2\lambda)$$



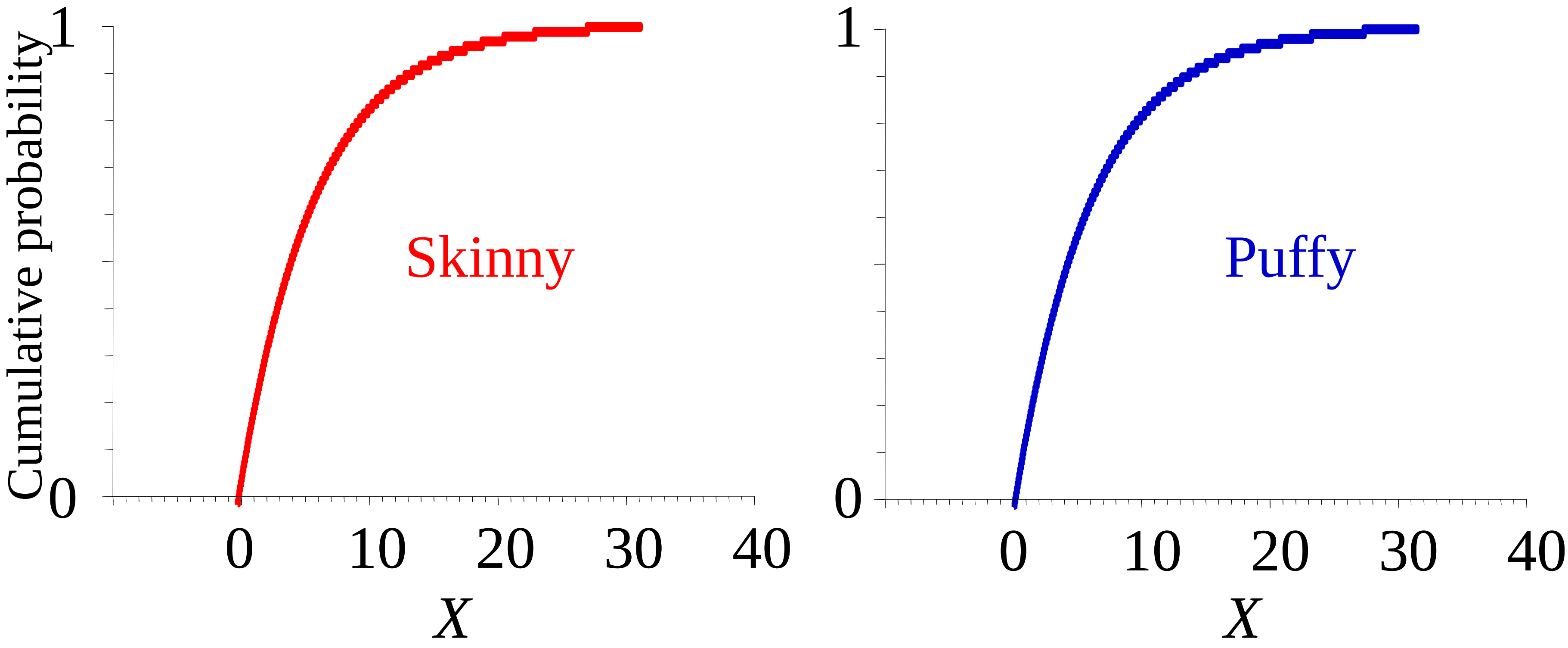
Maximum likelihood for multiple data



Results from uniforms approach



Negligible difference



Are these answers reasonable?

- Picking a single exponential to fit interval data
- Essentially no difference between Skinny and Puffy data sets, despite disparity of uncertainties
- Confidence bands actually smaller for Puffy, even though its uncertainty is much larger
- No guarantee that the answer approaches the true distribution *even if asymptotically many data are collected*

Take-home messages

- *Can* mix good and bad data in a consistent way
- ‘Uniforms’ approach more powerful if you’re comfortable with its assumptions
- Intervals approach weaker, but more reliable
- Data that would distinguish between the two approaches is exactly the data we don’t have

Further information: www.ramas.com/intstats.pdf

Computing with confidence

Many ways to fit distributions to data

- Maximum entropy
- Maximum likelihood
- Bayesian inference
- Method of matching moments ⟸ still most common
- Goodness of fit (KS, AD, χ^2, etc.)
- PERT
- Regression techniques
- Empirical distribution functions

…in fact there are even more methods…

Little coherence in practice

- Disparate methods used across risk analysis
- Common to mix and match distributions with different justifications
- Analyses are thus based on no clear criterion or standard of performance
- Is this okay?

Frequentist confidence intervals

- Favored by many engineers
- Guarantees statistical performance over time
- But difficult to employ consistently in analyses
- Not clear how to propagate them through mathematical calculations

Bayesian approaches

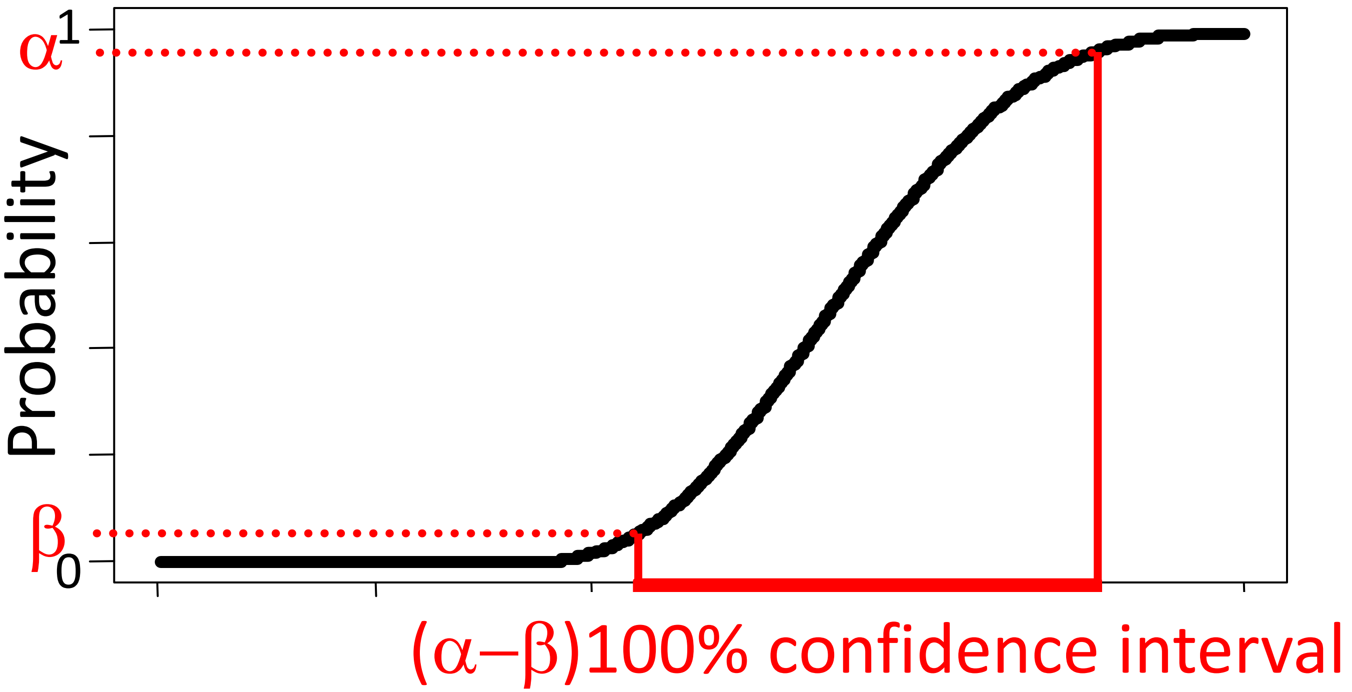
- Permit mathematical calculations
- But lack guarantees ensuring long-run statistical performance
- Many engineers are reluctant to use Bayesian methods

Confidence distributions

- Not widely used in statistics
- Introduced by Cox in the 1950s
- Closely related to well known ideas
  - Student's *t*-distribution
  - Bootstrap distributions

Confidence distributions

- Distributional estimators of (fixed) parameters
- Give confidence interval at *any* confidence level



Confidence interval

- A <u>confidence interval</u> with coverage α

  In replicate problems, a proportion α of computed confidence intervals will enclose the true value

- Using methods to compute confidence intervals thus ensures statistical *performance*

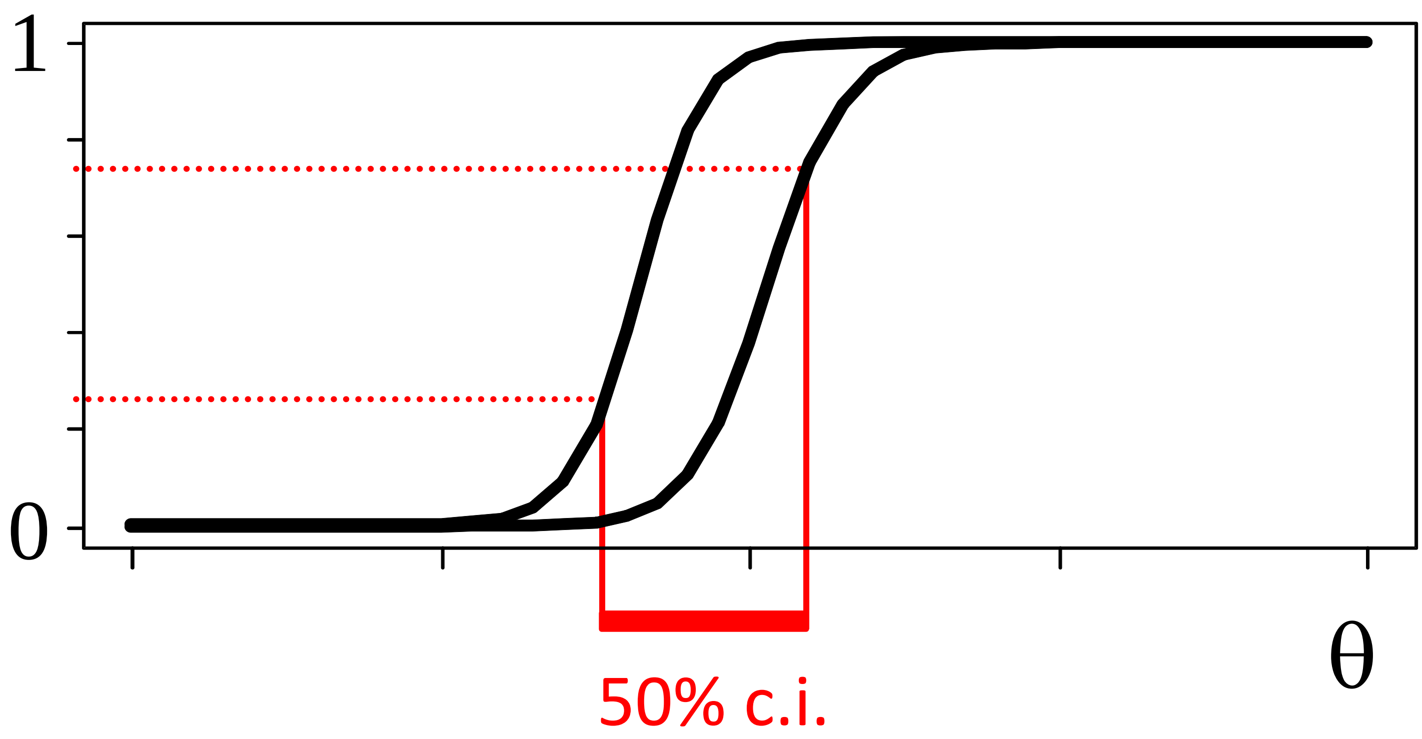
Confidence distributions

- Have the *shape* of a distribution
- But correspond to no random variables
- Not supposed to compute with them
- Don't always exist (e.g., for the binomial rate)

Confidence structures (c-boxes)

- Generalization of confidence distributions
- Reflect inferential uncertainty about parameter
- Known for many cases
  - binomial rate and other discrete parameters
  - normals, and many other problems
  - non-parametric case
- Still have performance/confidence interpretation

Confidence interpretation



Estimators

- Point estimates (e.g., sample mean)
- Interval estimates (e.g., confidence intervals)
- Distributional estimates (Bayesian posteriors)
- P-box estimates (e.g., c-boxes)

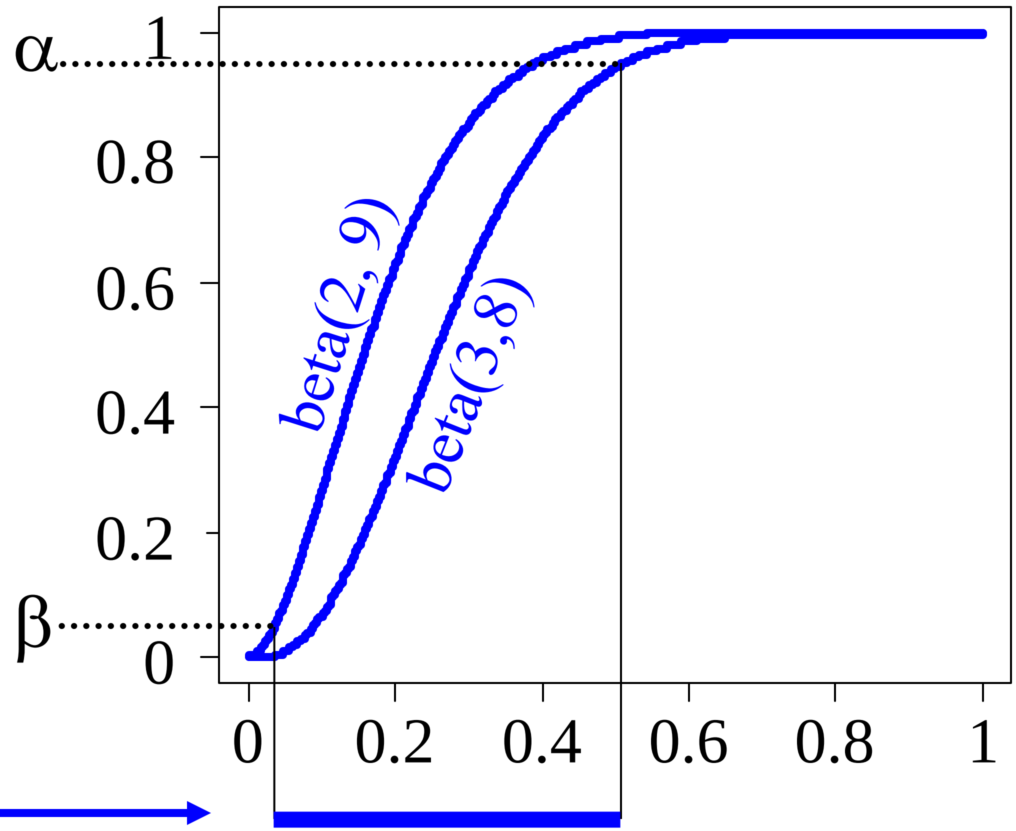
Binomial rate p for k of n trials

$p \sim \mathrm{env}(\mathrm{beta}(k,\, n{-}k{+}1),\, \mathrm{beta}(k{+}1,\, n{-}k))$

Data

$k = 2$

$n = 10$

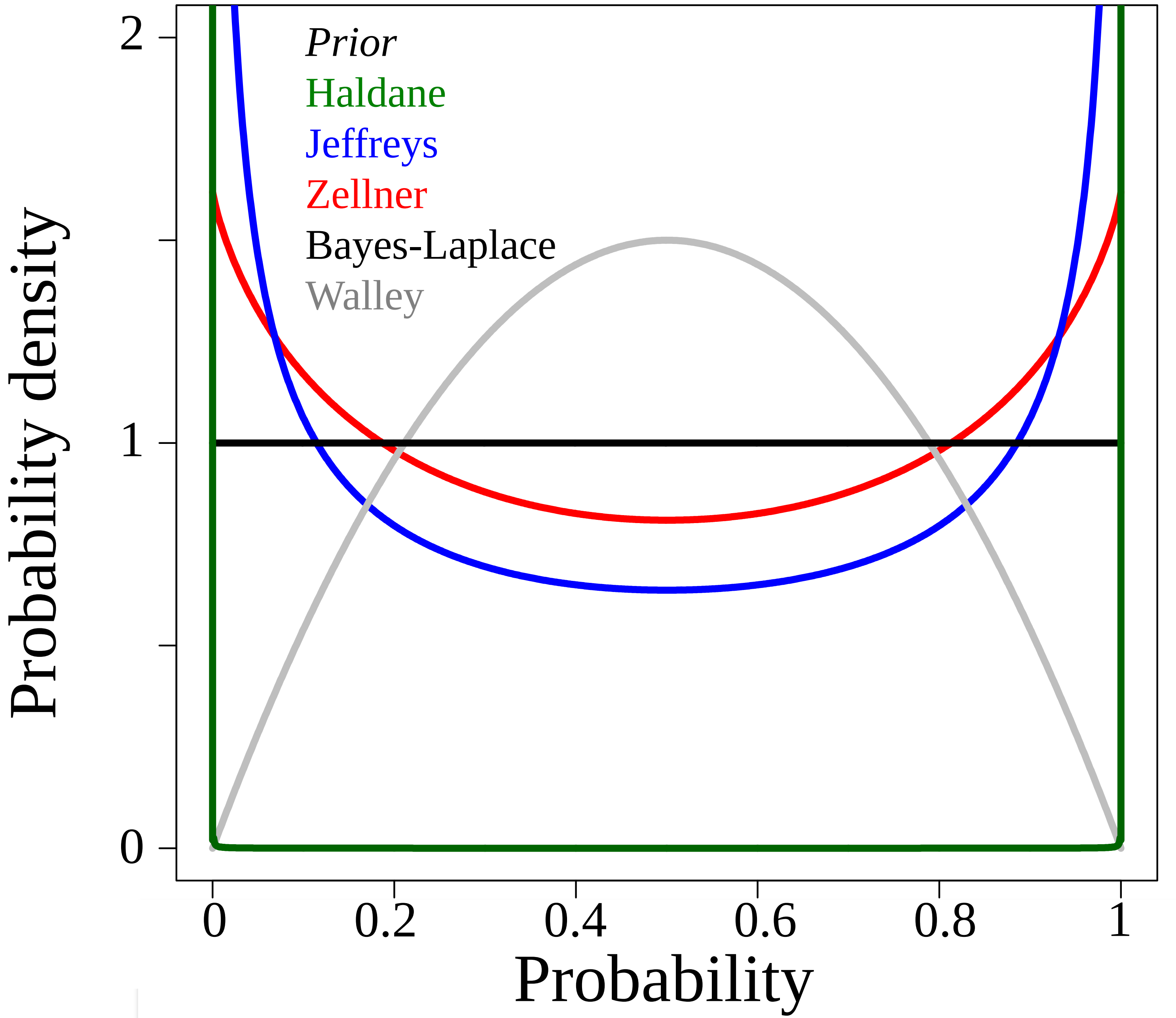


(α–β)100% confidence interval for p

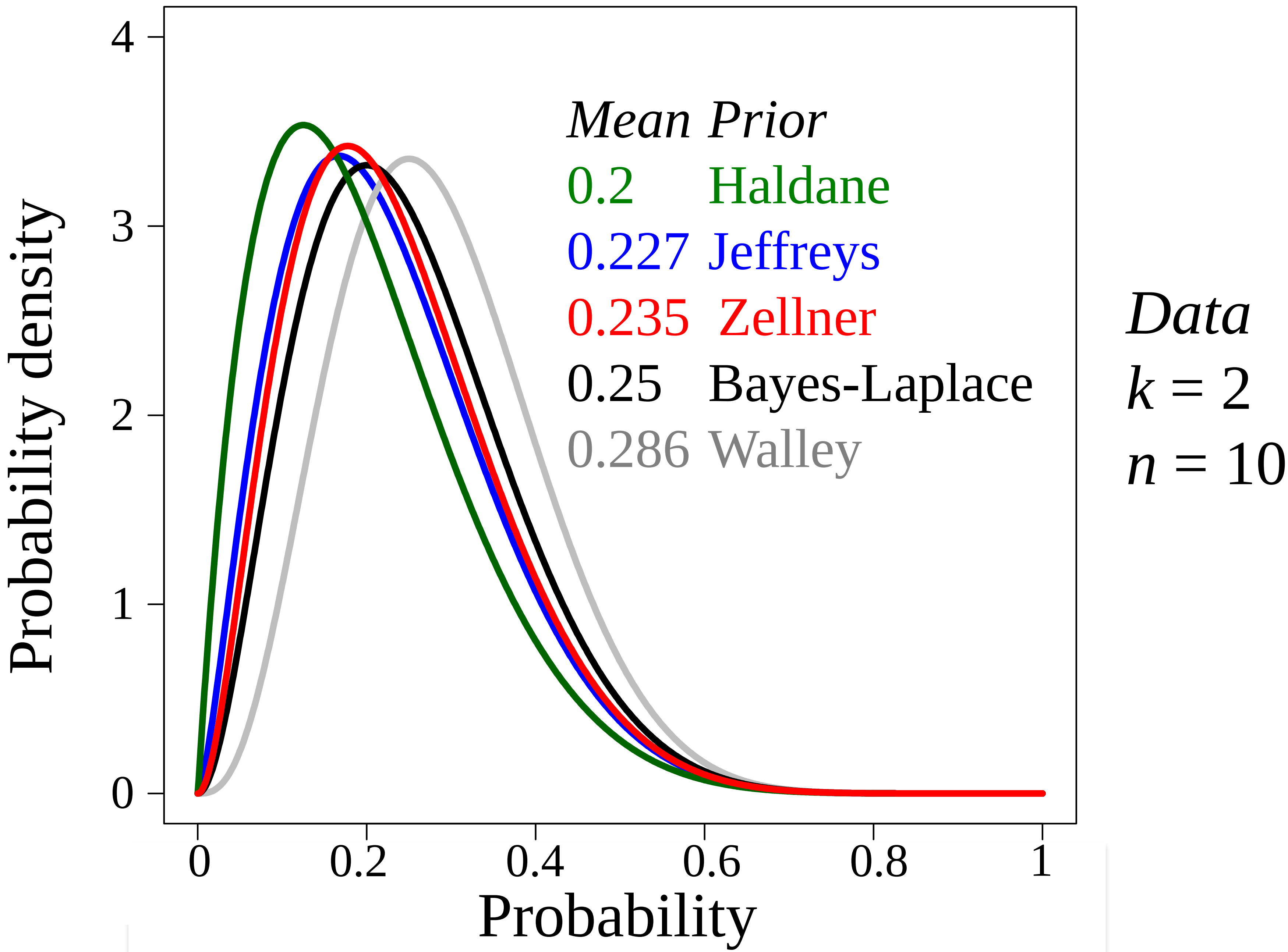
If 1–α = β, result is identical to classical Clopper–Pearson interval

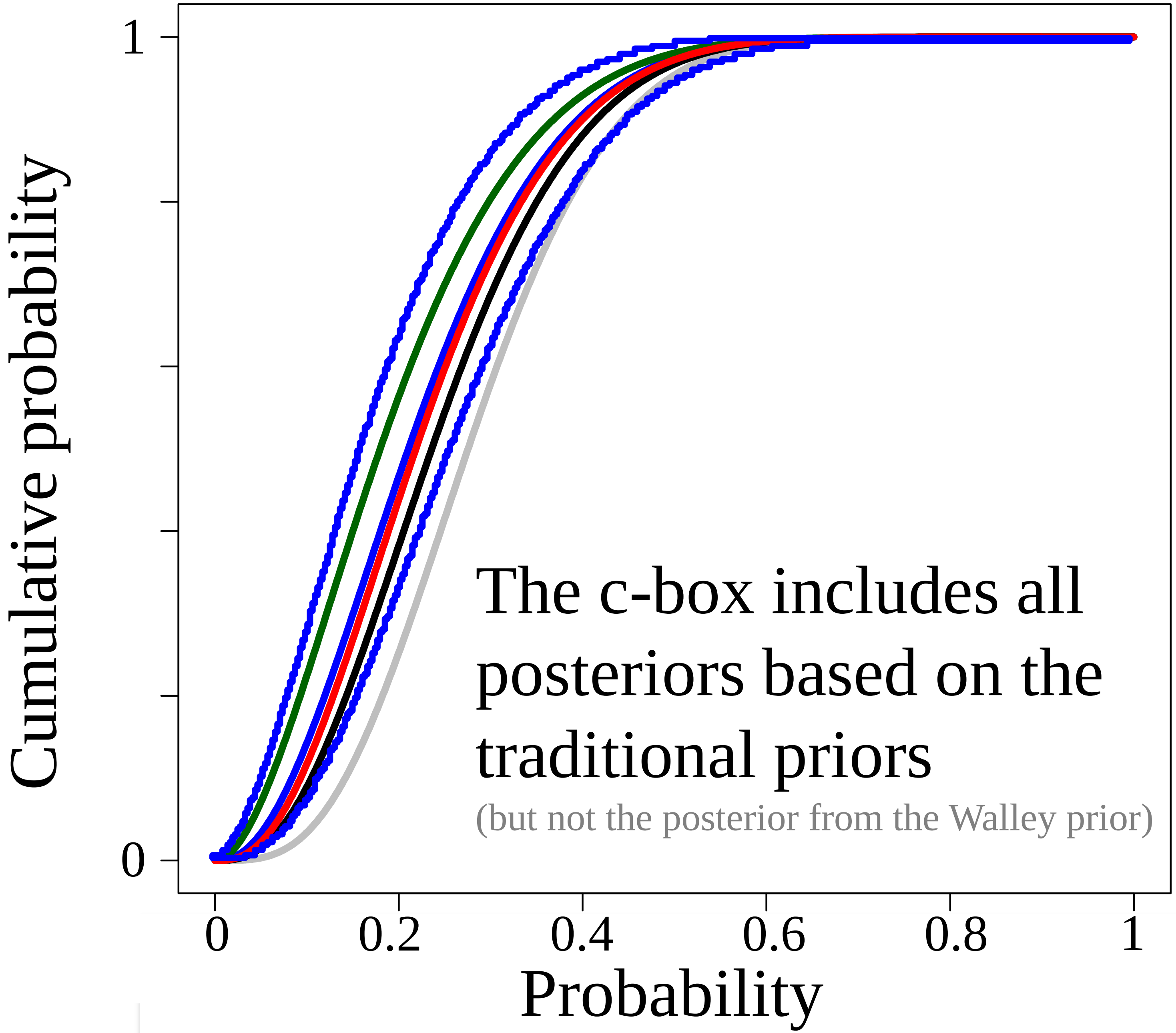
How does the Bayes analysis compare?

- No such thing as *the* Bayes analysis
- There are always many possible analyses
  - Different priors, which yield different answers
  - When data sets are small, the differences are big
- For binomial rate there are four or five priors Bayesians have not been able to chose among



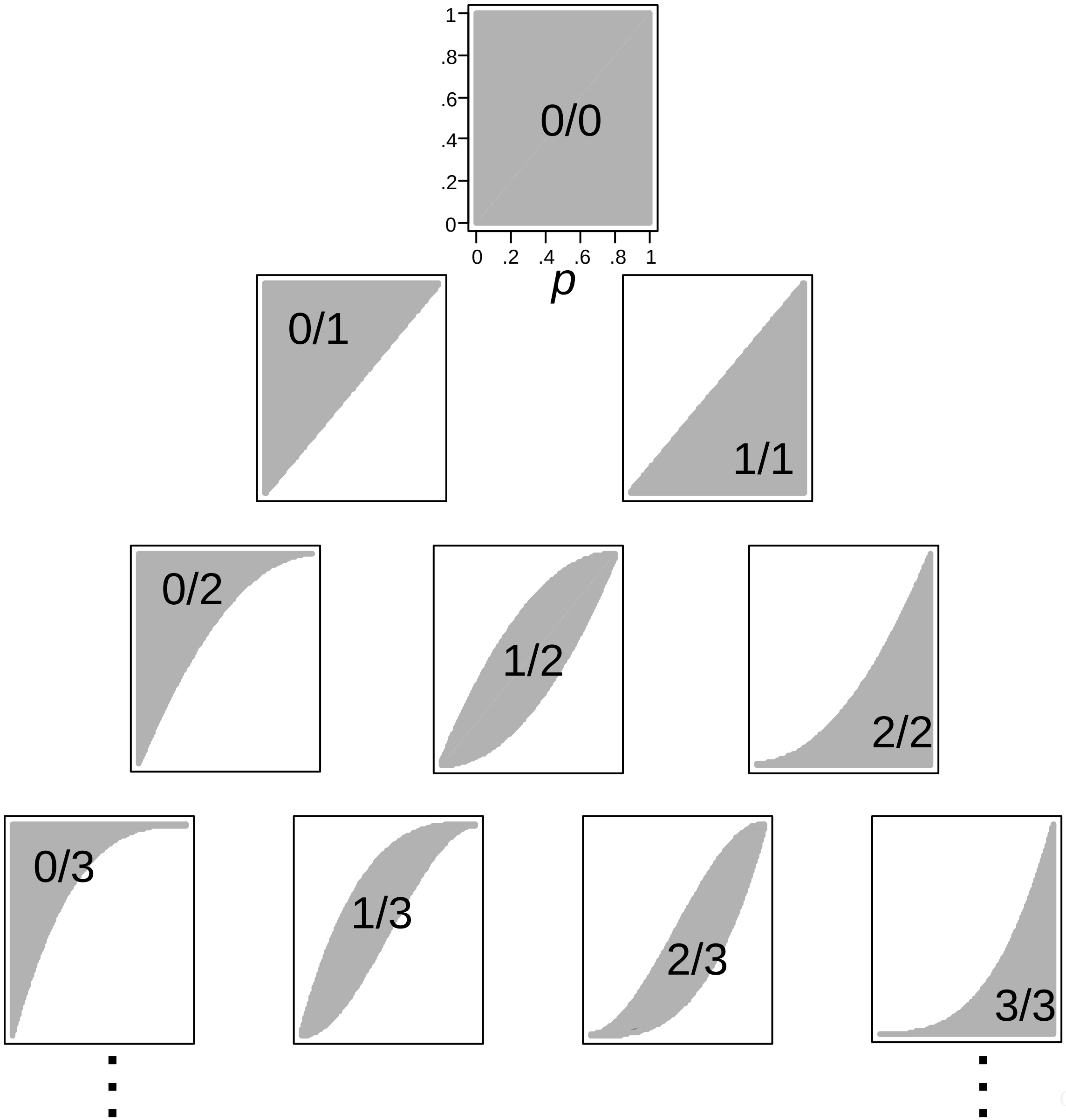
Prior
Haldane
Jeffreys
Zellner
Bayes-Laplace
Walley
Probability density
Probability
2
1
0
0
0.2
0.4
0.6
0.8
1


Probability density
4
3
2
1
0
0
0.2
0.4
0.6
0.8
1
Probability
Mean Prior
0.2 Haldane
0.227 Jeffreys
0.235 Zellner
0.25 Bayes-Laplace
0.286 Walley
Data
k = 2
n = 10



Cumulative probability
1
0
0
0.2
0.4
0.6
0.8
1
Probability
The c-box includes all posteriors based on the traditional priors
(but not the posterior from the Walley prior)

C-boxes partition the vacuous square

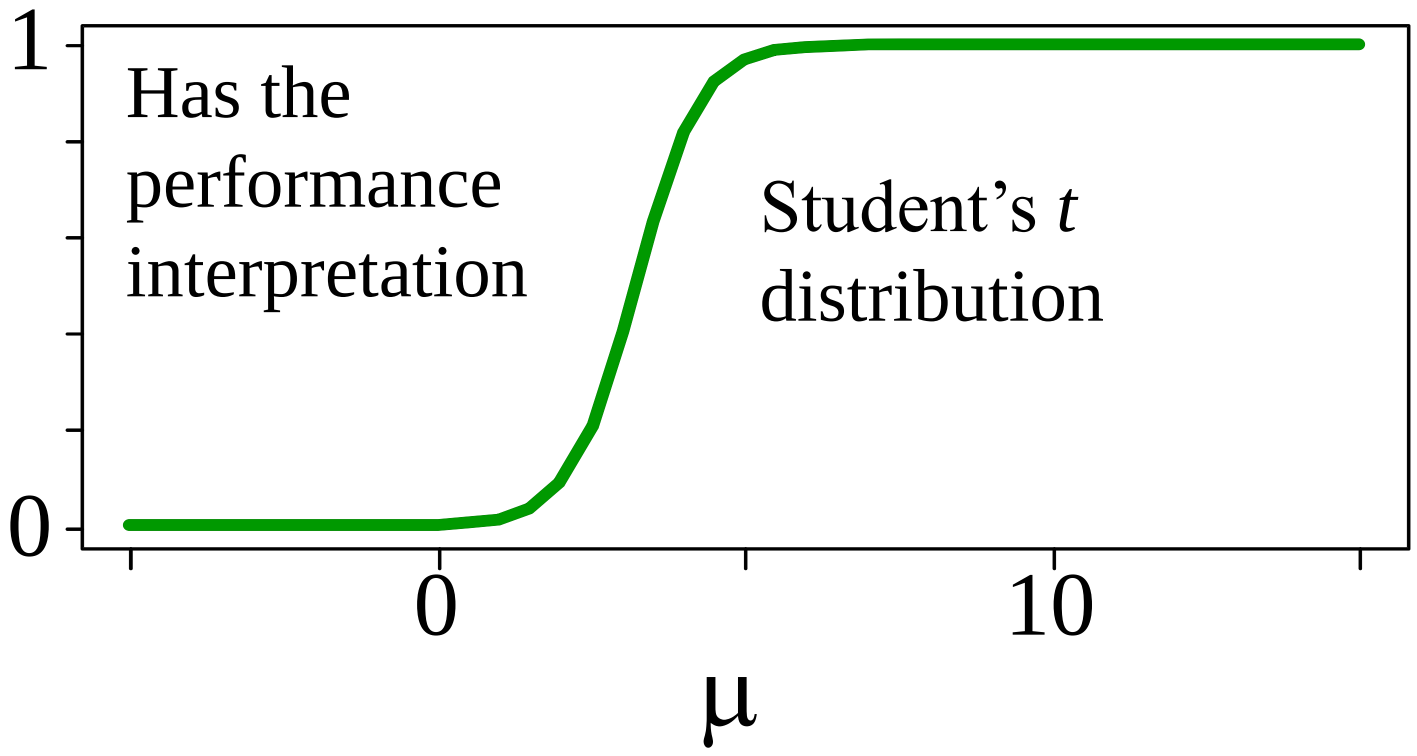


Example: normal mean

$$\mu \sim \bar{x} + s \cdot T_{n-1} / \sqrt{n}$$

Data

8
5.5
–1.3
3.5
0.8
2.8
1.8
2.2
3.5
5.3

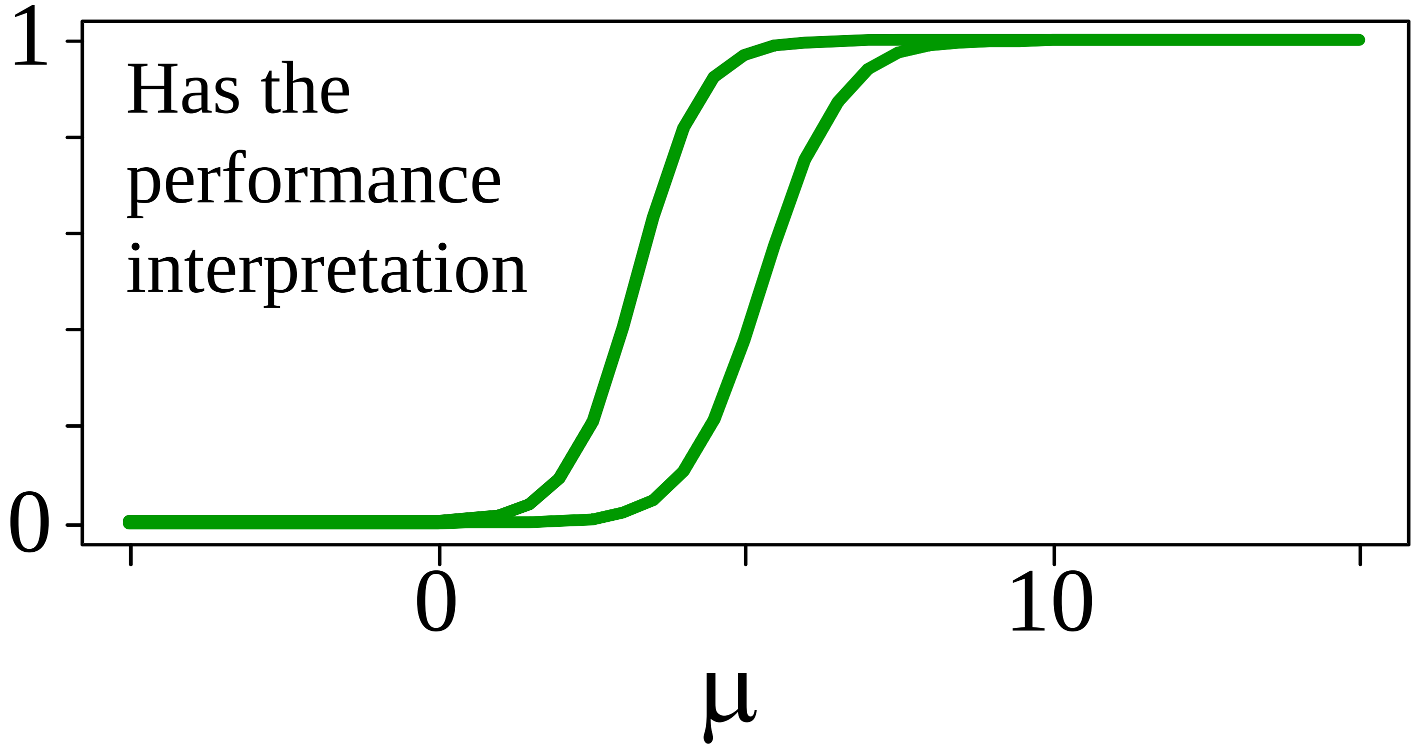


Example: normal mean

$$\mu \sim \bar{x} + s \cdot \mathrm{T}_{n-1} / \sqrt{n}$$

Data

[8,11]
[5.5,6.9]
[–1.3,0.3]
[3.5,7.5]
[0.8,1]
[2.8,4.2]
[1.8,5.2]
[2.2,5.2]
[3.5,5.7]
[5.3,6.1]



Deriving c-boxes

- Have to be derived for each special case
- Traditional approaches based on pivots
- Many solutions have been worked out

binomial(p, n), given n

binomial(p, n), given p

binomial(p, n)

Poisson(p)

⋮

normal(μ, σ)

lognormal(μ, σ)

gamma(a, b)

exponential(λ)

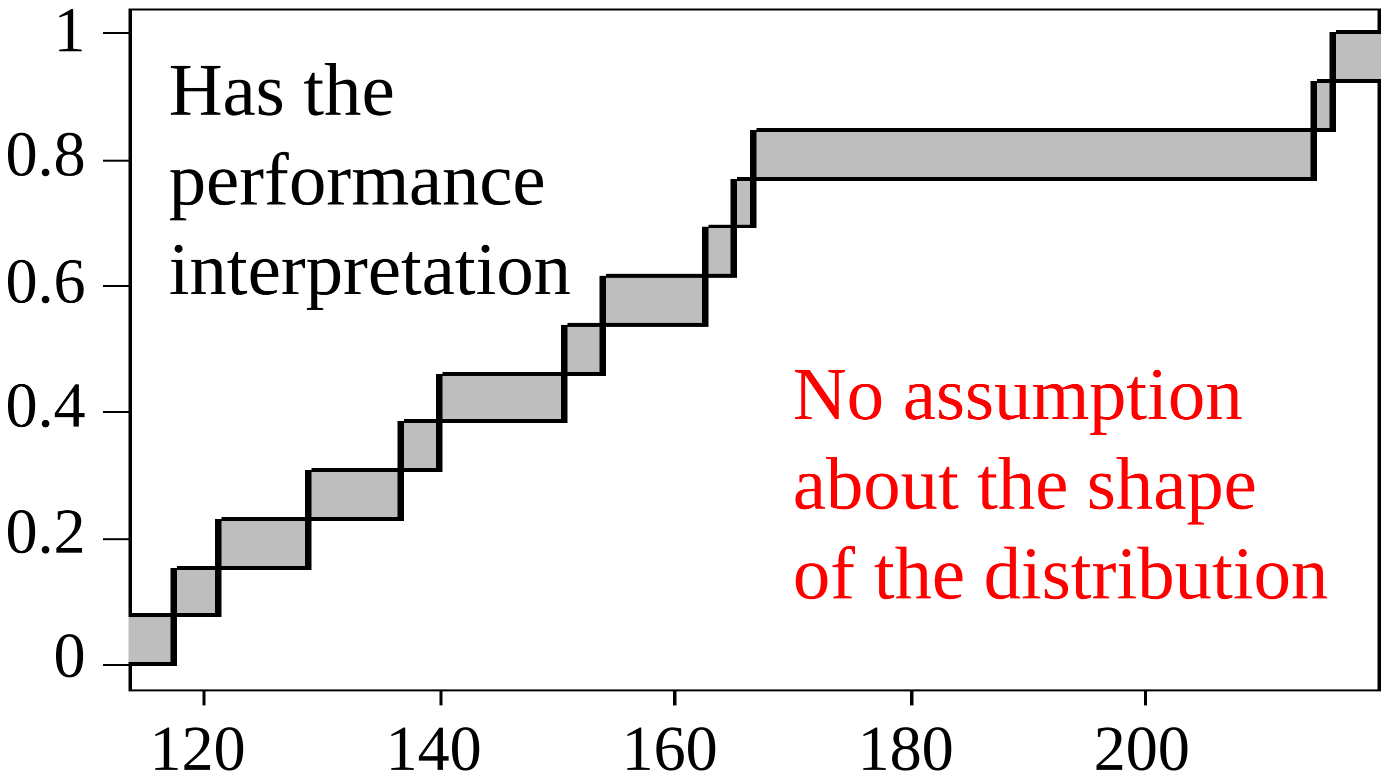
⋮

Example: non-parametric problem

$X \sim [(1+C(x))/(1+n),\ C(x)/(1+n)]$

where $C(x) = \#(X_i \leq x)$

Data
140.2
121.2
154
162.6
136.9
215.9
117.5
166.7
165.2
128.9
150.6
214.3



Captured uncertainties

- Uncertainty about distribution shape
- Sampling uncertainty (from small *n*)
- Measurement incertitude (±, censoring)
- Demographic stochsticity (integrality of data)

Propagated as probability boxes

- C-boxes can be combined in mathematical expressions using the p-box technology

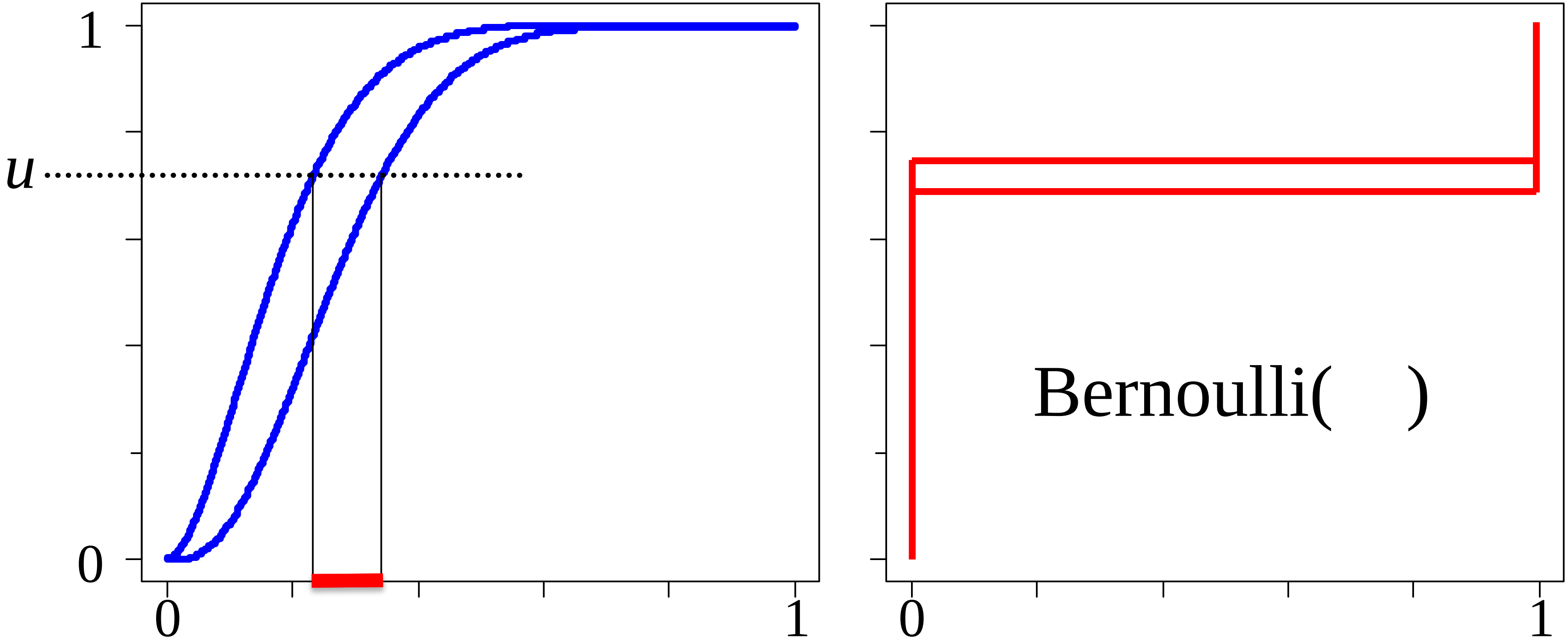
- *Results also have performance interpretations*

- C-boxes can also make predictive p-boxes
  - Analogous to frequentist prediction distributions
  - Or Bayesian posterior predictive distributions

Prediction structures

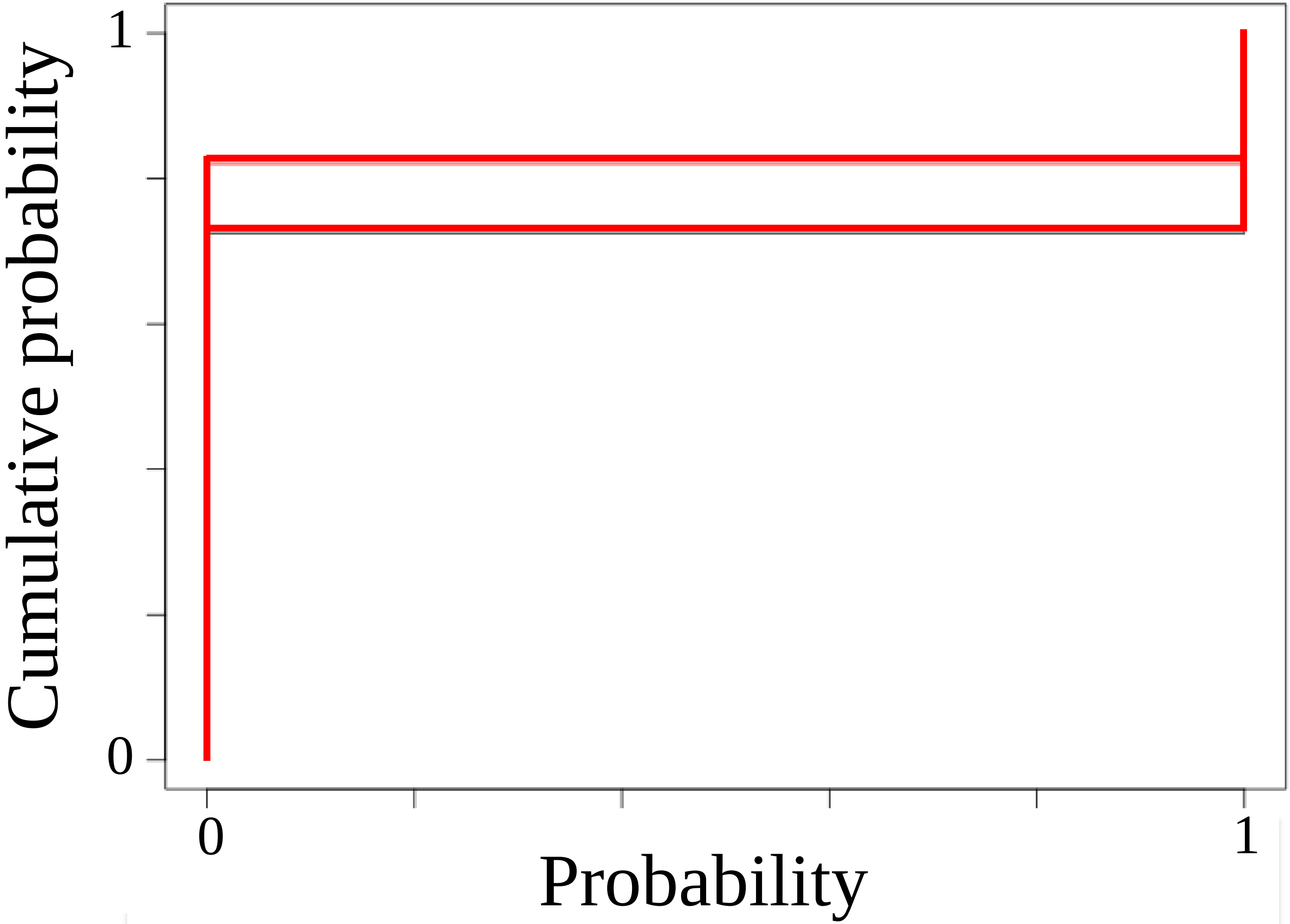
- C-boxes can model the uncertainty about the *underlying distribution* that generated the data
- Stochastic mixture of p-boxes from interval parameters specified by slices from the c-box
- This is a *composition* of the c-box through the probability model

Example: Bernoulli distribution



Each interval slice defines a p-box for the underlying distribution (rather than a precise distribution)

Average all such p-boxes



Predictive p-boxes for *observables*

Bernoulli trials $b_i \sim \mathrm{B}(p)$, $i=1,\ldots,n$

$b_{n+1} \sim \mathrm{Bernoulli}([k, k+1]/(n+1))$, $k = \Sigma_i b_i$

Random binomial sample $k \sim \mathrm{binomial}(p, n)$

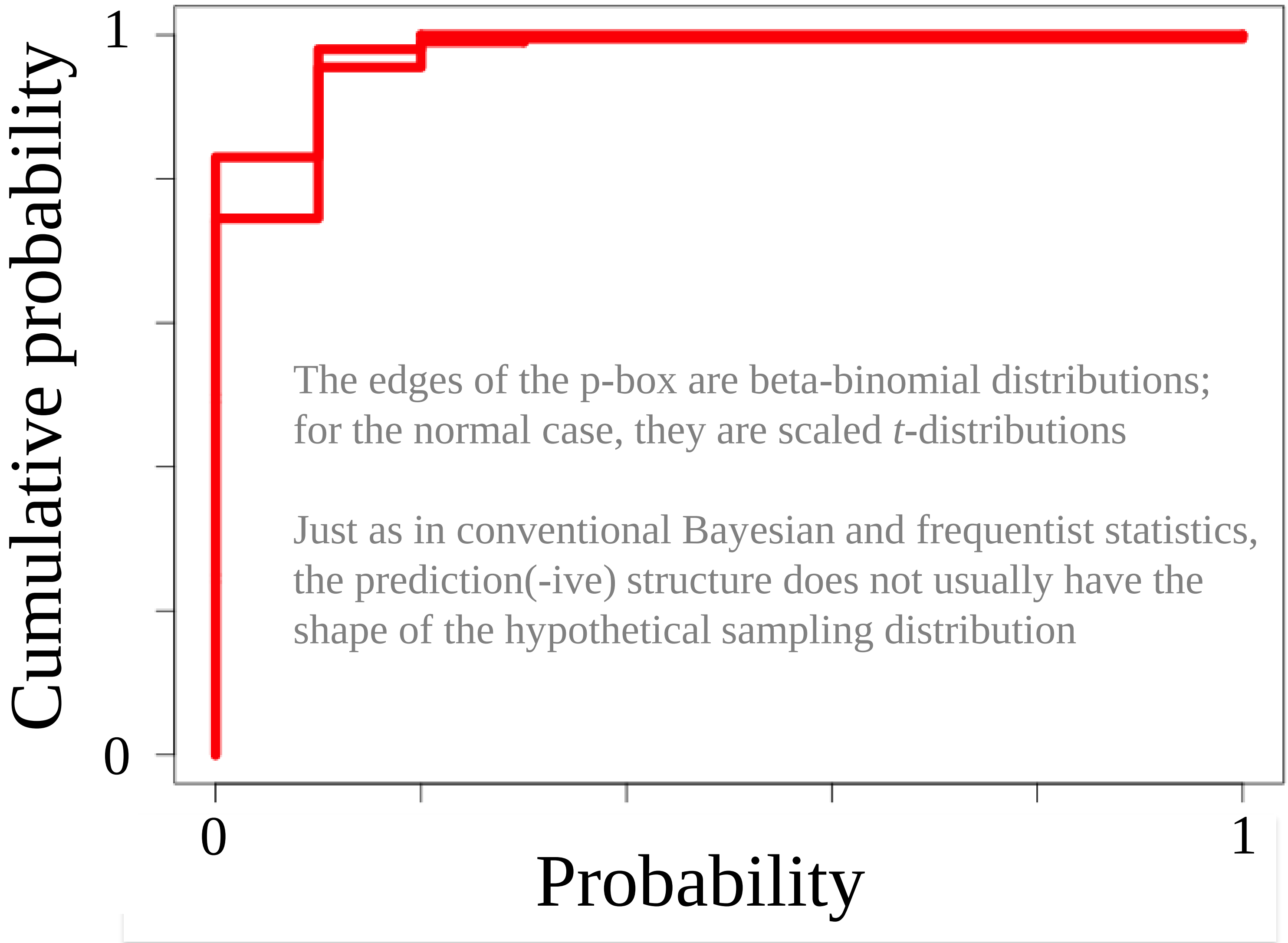
$k_2 \sim [\mathrm{BB}(k, nN-k+1, N), \mathrm{BB}(k+1, nN-k, N)]$

Random normal samples $X_i \sim \mathrm{N}(\mu, \sigma)$, $i=1,\ldots,n$

$X_{n+1} \sim \mathrm{mean}(X_i) + \mathrm{stdev}(X_i) \cdot \mathrm{sqrt}(1+1/n) \cdot \mathrm{T}(n-1)$

BB denotes beta-binomial; T denotes Student's *t*

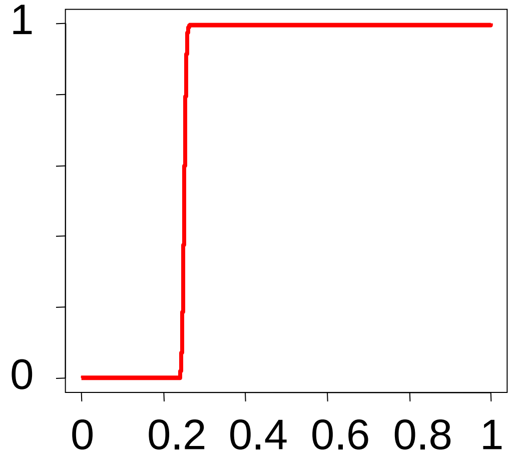
Example: Binomial distribution



Prediction structures are p-boxes

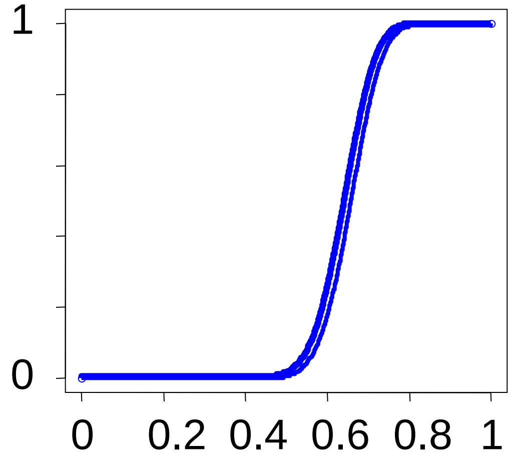
- Results also have the confidence interpretation, but the intervals are *prediction intervals* rather than confidence intervals
- Prediction intervals enclose a specified percentage of observable values, on average
- It is also possible to derive analogous tolerance structures, which encode *tolerance intervals* (X% sure to enclose Y% of the population)

Computing with c-boxes



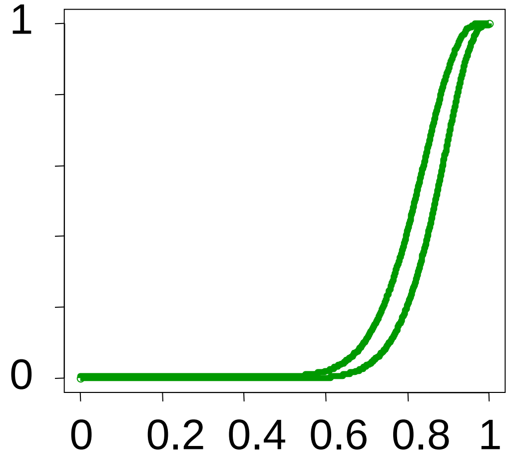
Plan A

25% fail



Plan B

39 out of 60 failed

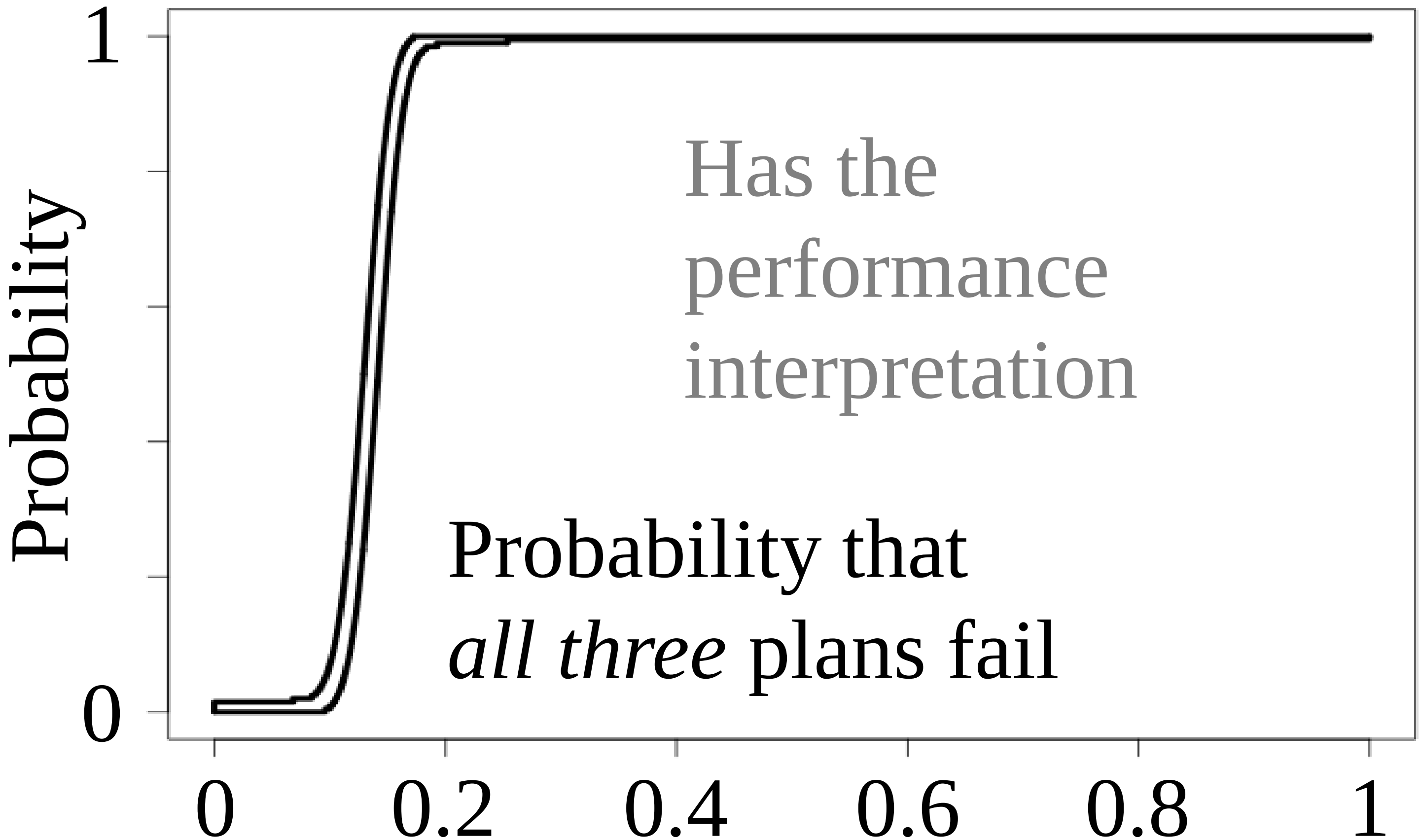


Plan C

17 out of 20 failed

What if we used all three plans independently?

Conjunction (AND)



Summary for c-boxes

- Confidence boxes carry inferential uncertainties through mathematical operations
- Give confidence intervals on results at any α level
- Defined by performance, so not unique
  - Just as confidence intervals are not unique
  - May create some flexibility
- Don't seem to be overly conservative
  - Elaborate simulation studies have so far not found this

Applies even with zero data

- There may be no *sample* data at all
- If constraints are known that specify a rigorous p-box, then it encodes prediction intervals
- So our performance interpretation applies for
  - Parametric problems
  - Nonparametric problems
  - No data problems

Maximum likelihood

Maximum entropy

Bayesian inference

Estimation

Expert elicitation

“the great frontier of making things up”

Method of moments

PERT

Conclusions

- C-boxes characterize risk analysis inputs given limited sample or constraint information
- Reasonable answers when data and tenable assumptions don’t justify particular distributions
- C-boxes don’t optimize; they *perform*
- C-boxes could serve as the lexicon in a language of risk analysis

More information

https://sites.google.com/site/confidenceboxes/

- Papers
- Slide presentations
- Free software

Google “confidence boxes” [plural…singular is a blog on teenage self-esteem & self-empowerment]

C-boxes are Bayesian

Bayesian sensitivity analysis

- Under robust Bayes approach, c-boxes can be thought of as Bayesian posteriors
- Don't require specification of a unique prior
- Have added feature of statistical performance
- Imply posterior predictive distributions
- Compatible with specifying a robust or precise prior when that's desirable

Prediction structures are p-boxes

- Also have confidence interpretation
  - Results are *prediction intervals* enclosing specified percentage of observable values, on average

- Can also define analogous tolerance structures
  - *Tolerance intervals* are *X*% sure to enclose *Y*% of the population

Fermi estimates

No data at all

Fermi estimates

http://en.wikipedia.org/wiki/Fermi_problem

- Enrico Fermi
  - Estimated strength of Trinity Test bomb by dropping scraps of paper from his hand during the blast and seeing how they fell
  - He also estimated the number of piano tuners in Chicago

5 million people in Chicago
On average, 2 people / home
Roughly 1 home in 20 has a piano that's tuned regularly
Regular tuning happens about once per year
(5e6 people)/(2 people/home)×(1 piano/20 home)×(1 tuning/piano/yr)=125,000 tunings/yr
Tuning a piano takes about two hours
A tuner works 8 hours a day, 5 days a week, and 50 weeks a year
(50 wk/yr)×(5 d/wk)×(8 hr/d)/(1 tuning per 2 hr per tuner) = 1000 tunings per yr per tuner
(125,000 tunings per year) / (1000 tunings per year per piano tuner) = 125 piano tuners

Simplicity has advantages

- Clarity
  - Dimensions and units analysis
  - Approximation may be good enough (often within order of magnitude)
- Good to do a cheap calculation first
  - People invest too much in complex calculation
  - The complexity can obscure a large error
  - Simplicity lets you *see* the assumptions and check that the calculations make sense
  - Tells us how to improve the calculation

Probabilizing Fermi

- Seat-of-the-pants methods to estimate distributions and correlations

- Probabilistic convolution (MC simulation)

- What you need:
  - Model
  - Marginal distributions
  - Correlations
  - Assumptions that the above are good

With these, and my dad's barn, we can put on a risk analysis!

Seat-of-the-pants distributions

- Just *assume* shape, e.g., lognormal
- *Guesstimate* parameters crudely
  - Ask experts the range of possible values
  - Take the geometric mean of upper and lower

$$\mathrm{GM} = \sqrt{X_U X_L}$$

  - Ballpark g.s.d. by guessing the range's coverage

Geometric Standard Deviation

$$\mathrm{GSD} = (X_U / X_L)^{1/(2z_p)}$$

z_p is the standard normal deviate at probability p

Bowman, J.D., S.A. Shulman, and S. Sivagenesan. 2010. Expanded Fermi estimates for variances and correlations from expert judgments. [manuscript and slide presentation]. NIOSH and University of Cincinnati.
Pelega, M., M.D. Normanda, J. Horowitzb and M.G. Corradinia. 2007. An expanded Fermi solution for microbial risk assessment. *International Journal of Food Microbiology* 113(1): 92-101. http://www.ncbi.nlm.nih.gov/pubmed/17014921

R functions for Fermi (log)normal

```
fermi.lnorm = function(x1, x2, pr=0.9) {
  gm = sqrt(x1*x2)
  gsd = sqrt(x2/x1) ^ (1/qnorm(pr))
  log(c(gm, gsd))
  }
```

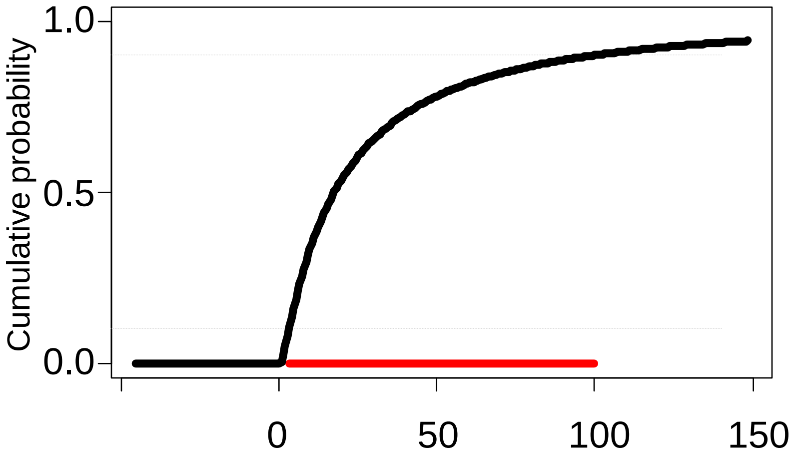
lognormal

```
fermi.dlnorm = function(x, meanlog = 0, sdlog = 1, log = FALSE) {t = fermi.lnorm(x1,x2,pr); dlnorm(x,t
fermi.plnorm = function(q, meanlog = 0, sdlog = 1, lower.tail = TRUE, log.p = FALSE) {t = fermi.lnorm(
fermi.qlnorm = function(p, meanlog = 0, sdlog = 1, lower.tail = TRUE, log.p = FALSE) {t = fermi.lnorm(
fermi.rlnorm = function(n, meanlog = 0, sdlog = 1) {t = fermi.lnorm(x1,x2,pr); rlnorm(n,t[[1]],t[[2]])}

fermi.norm = function(x1, x2, pr=0.9) {
  m = (x1 + x2) / 2
  s = (x2 - x1) / (2 * qnorm(pr))
  c(m, s)
  }
```

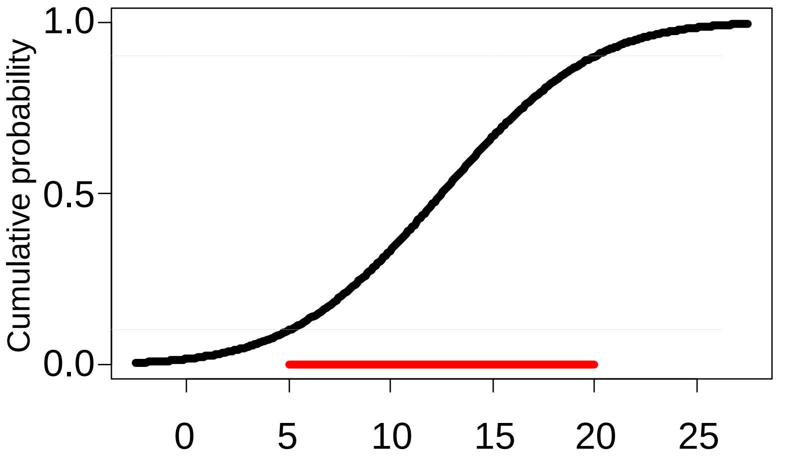
normal

```
fermi.dnorm = function(x, x1, x2, pr=0.9, log = FALSE) {t = fermi.norm(x1,x2,pr); dnorm(x,t[[1]],t[[2]],
fermi.pnorm = function(q, x1, x2, pr=0.9, lower.tail = TRUE, log.p = FALSE) {t = fermi.norm(x1,x2,pr);
fermi.qnorm = function(p, x1, x2, pr=0.9, lower.tail = TRUE, log.p = FALSE) {t = fermi.norm(x1,x2,pr);
fermi.rnorm = function(n, x1, x2, pr=0.9) {t = fermi.norm(x1,x2,pr); rnorm(n,t[[1]],t[[2]])}
```

Fermi examples



> fermi.lnorm(3,100)
[1] 2.851891 1.368091



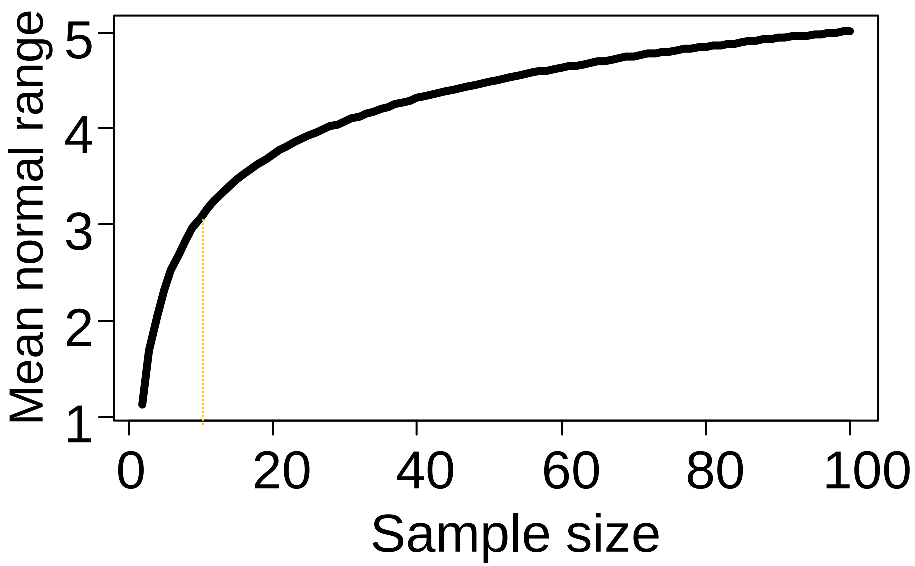
> fermi.norm(5,20)
[1] 12.500000 5.852281

Distribution shape is known
Coverage *p* value is known
Range is symmetric
No sampling uncertainty

Mean range trick

- Alternative approach to estimate standard deviation
- Divide observed range by mean range for the sample size
- Assumes
  - random samples
  - normality

If the range of 10 values is [52,86]
$s \approx 34/3 \approx 11$



Gumbel, E.J. 1958. *Statistics of Extremes*, Columbia University Press (§ 8.3.4)

R functions for mean range trick

```
mnr = function(n,many=10000) {
  xL = xU = rnorm(many)
  for (i in 2:n) {
    xx = rnorm(many)
    xL = pmin(xL,xx)
    xU = pmax(xU,xx) }
  mean(xU - xL)
  }
```

lognormal

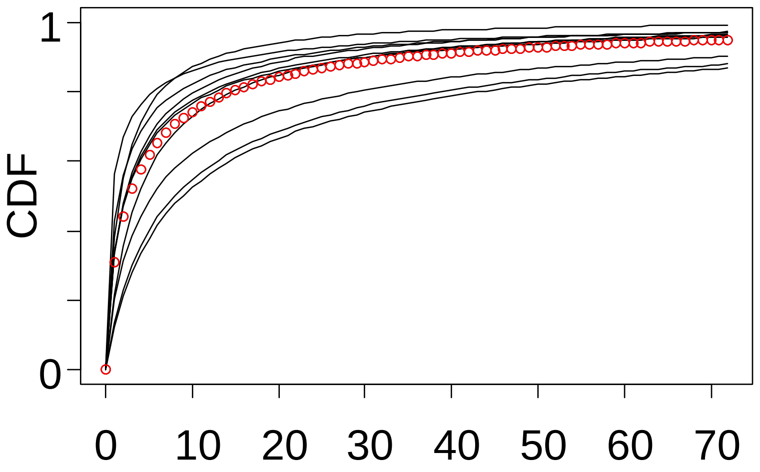
```
fermi.lnorm = function(x1, x2, n=NULL, pr=0.9) {
  gm = sqrt(x1*x2)
  if (is.null(n)) gsd = sqrt(x2/x1) ^ (1/qnorm(pr)) else gsd = exp((log(x2) - log(x1)) / mnr(n))
  log(c(gm, gsd))
  }
```

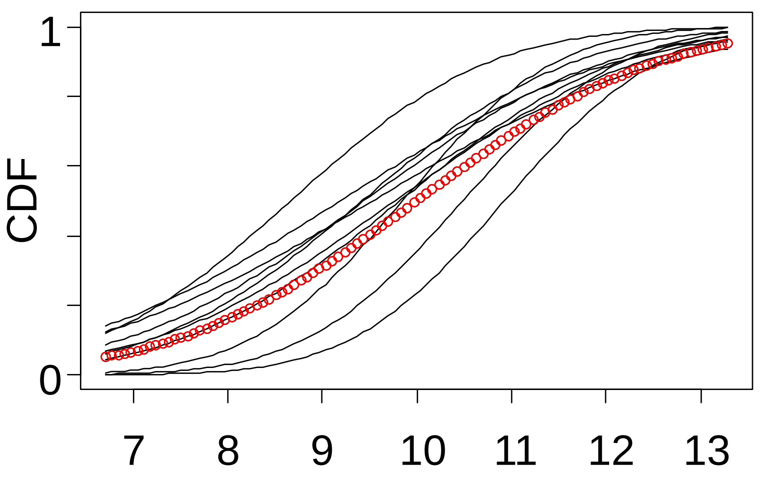
normal

```
fermi.norm = function(x1, x2, n=NULL, pr=0.9) {
  m = (x1 + x2) / 2
  if (is.null(n)) s = (x2 - x1) / (2 * qnorm(pr)) else s = (x2 - x1) / mnr(n)
  c(m, s)
  }
```

Examples with mean range

- True distribution in red
- Sample 20 points
- Use mean range trick
- Show result in black





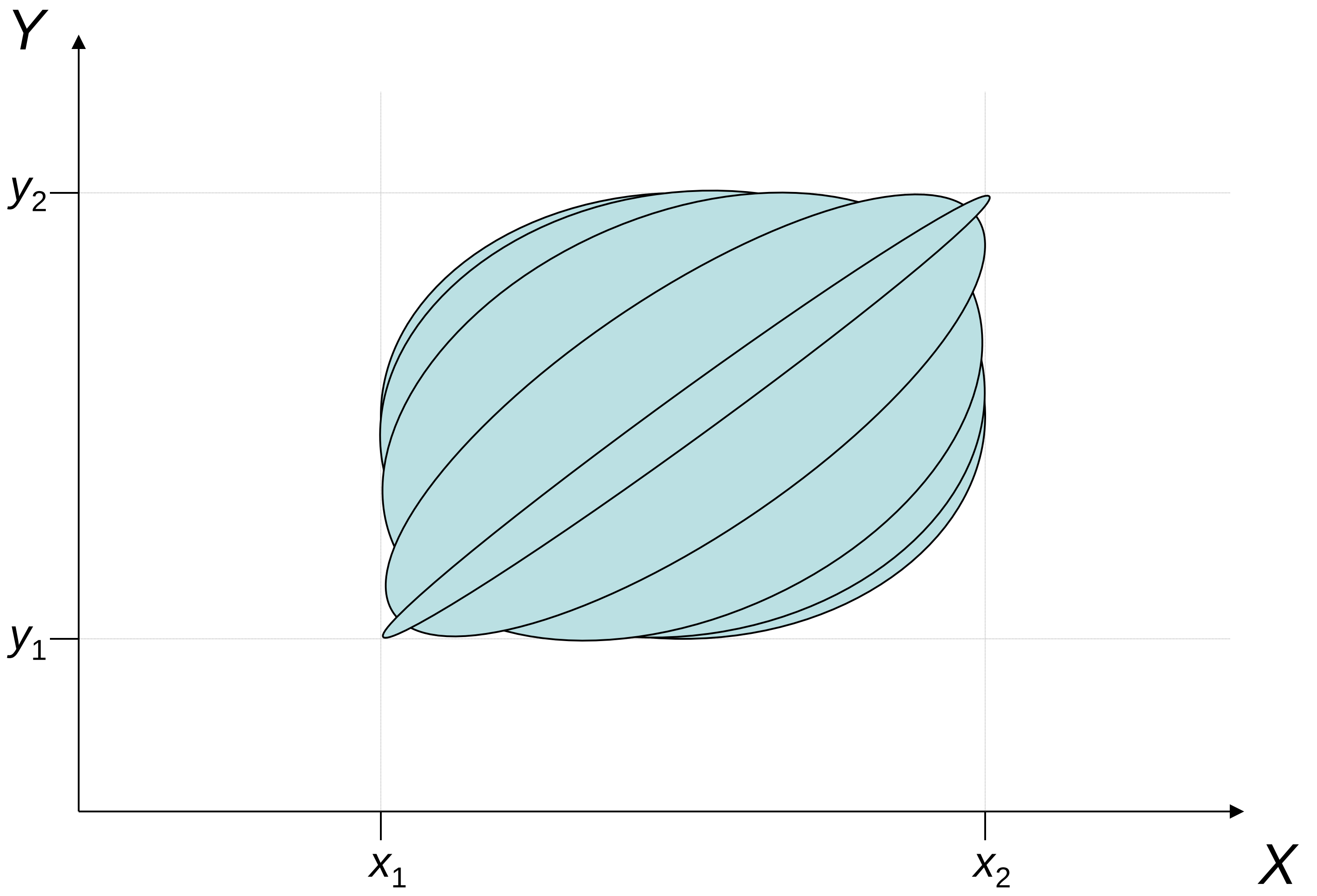
Fermi strategies

- Work with many distribution shapes
  - Not clear it'll work with 3-parameter shapes
- Can pool input from several experts
  - Not clear how to weight their contributions
- Make many assumptions
  - Distribution shape is known
  - Sample values not accessible, but range is
  - Can specify coverage, or sample size
  - Symmetric coverage, or random sampling

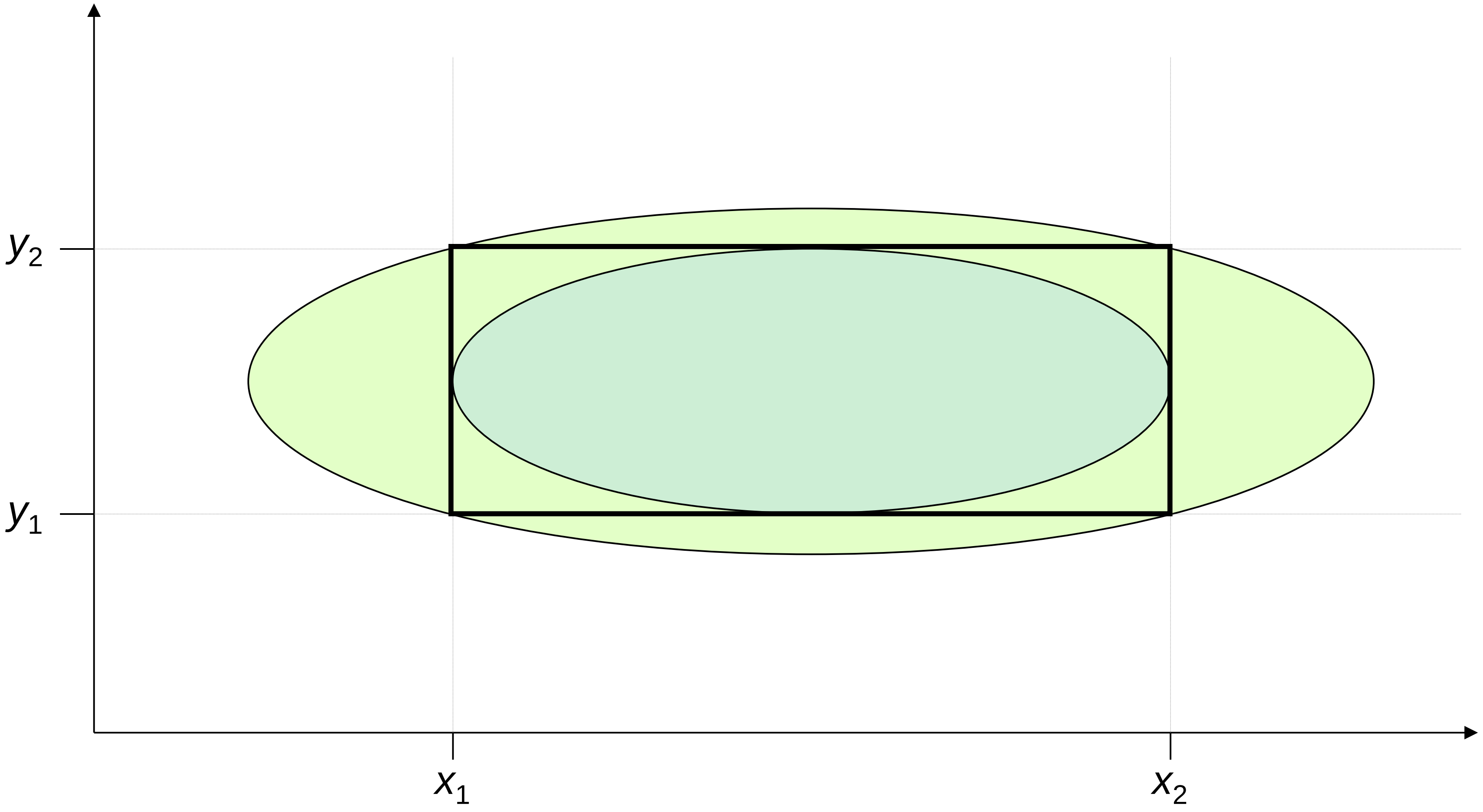
What about correlations?

- Could just ask the experts about them too
  - This approach yields very poor results
- Could try to get from data
  - Much harder than getting marginal distributions
- Could use a graphical approach

Correlation within ellipses



Two-dimensional ranges

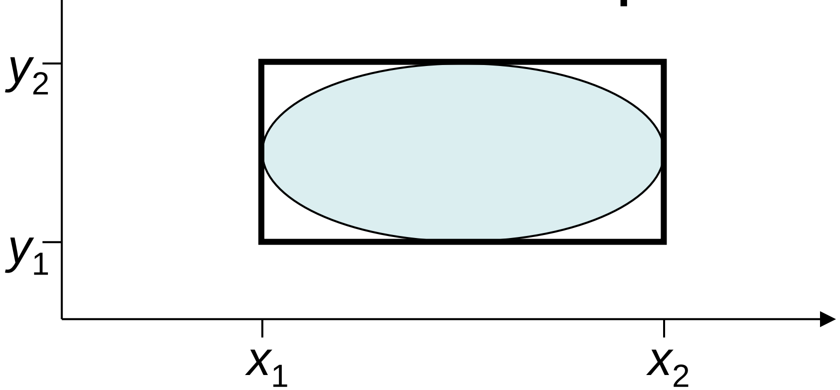


Ellipses as multivariate intervals

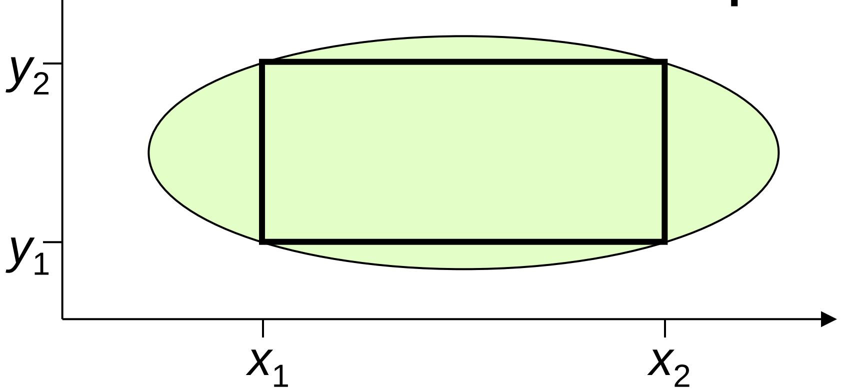
- Use the inscribed ellipse if the joint extremes within the box are impossible
- Use the circumscribed ellipse if they are possible and you want to be conservative

Formulas for the ellipses

Inscribed ellipse



Circumscribed ellipse



$$\frac{1}{\Delta_x^2}\left(x-\tilde{x}\right)^2-\frac{1}{\Delta_y^2}\left(y-\tilde{y}\right)^2=1$$

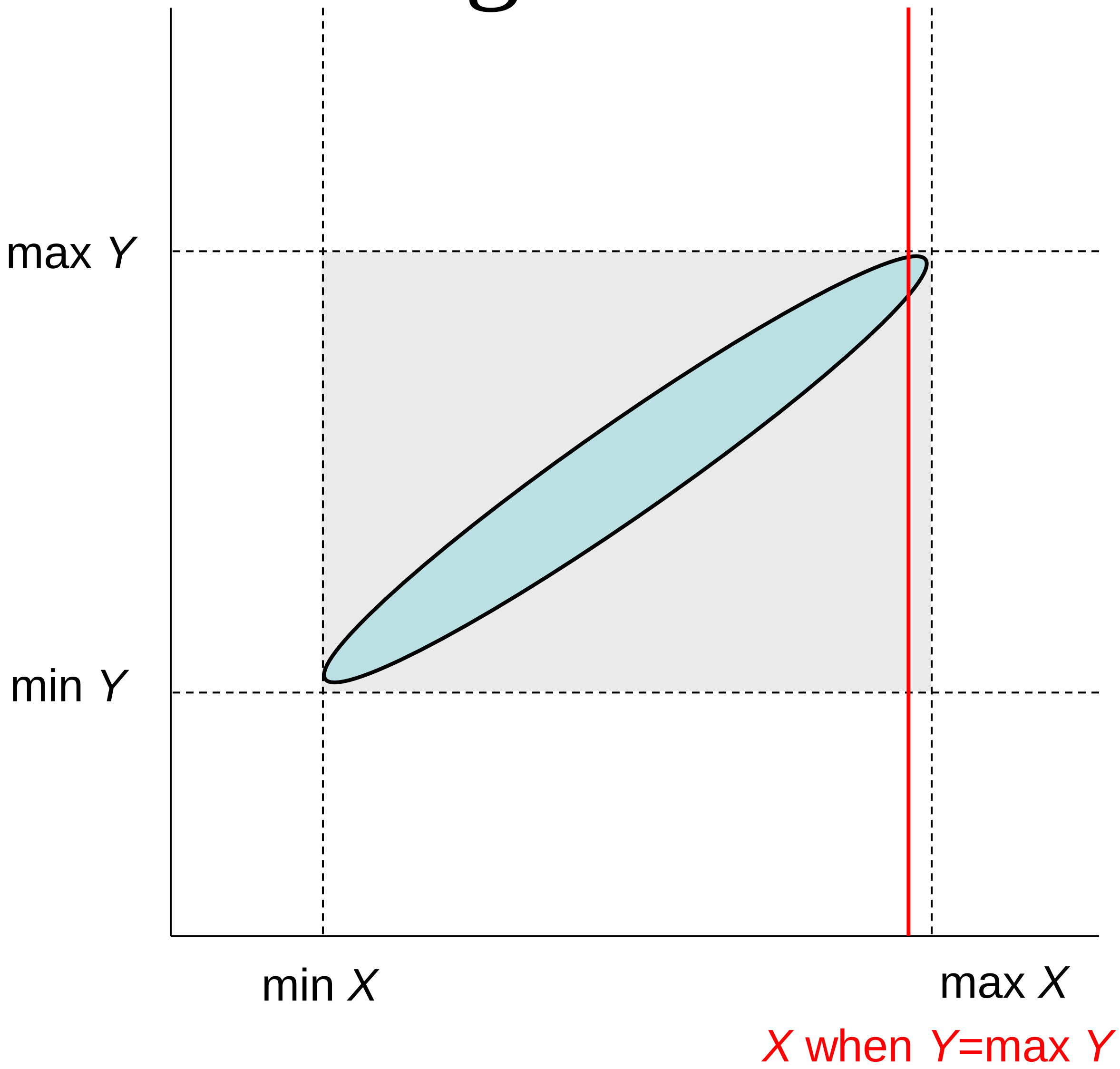
$$\Delta_x=\frac{x_2-x_1}{2} \qquad \Delta_y=\frac{y_2-y_1}{2}$$

$$\tilde{x}=\frac{x_1+x_2}{2} \qquad \tilde{y}=\frac{y_1+y_2}{2}$$

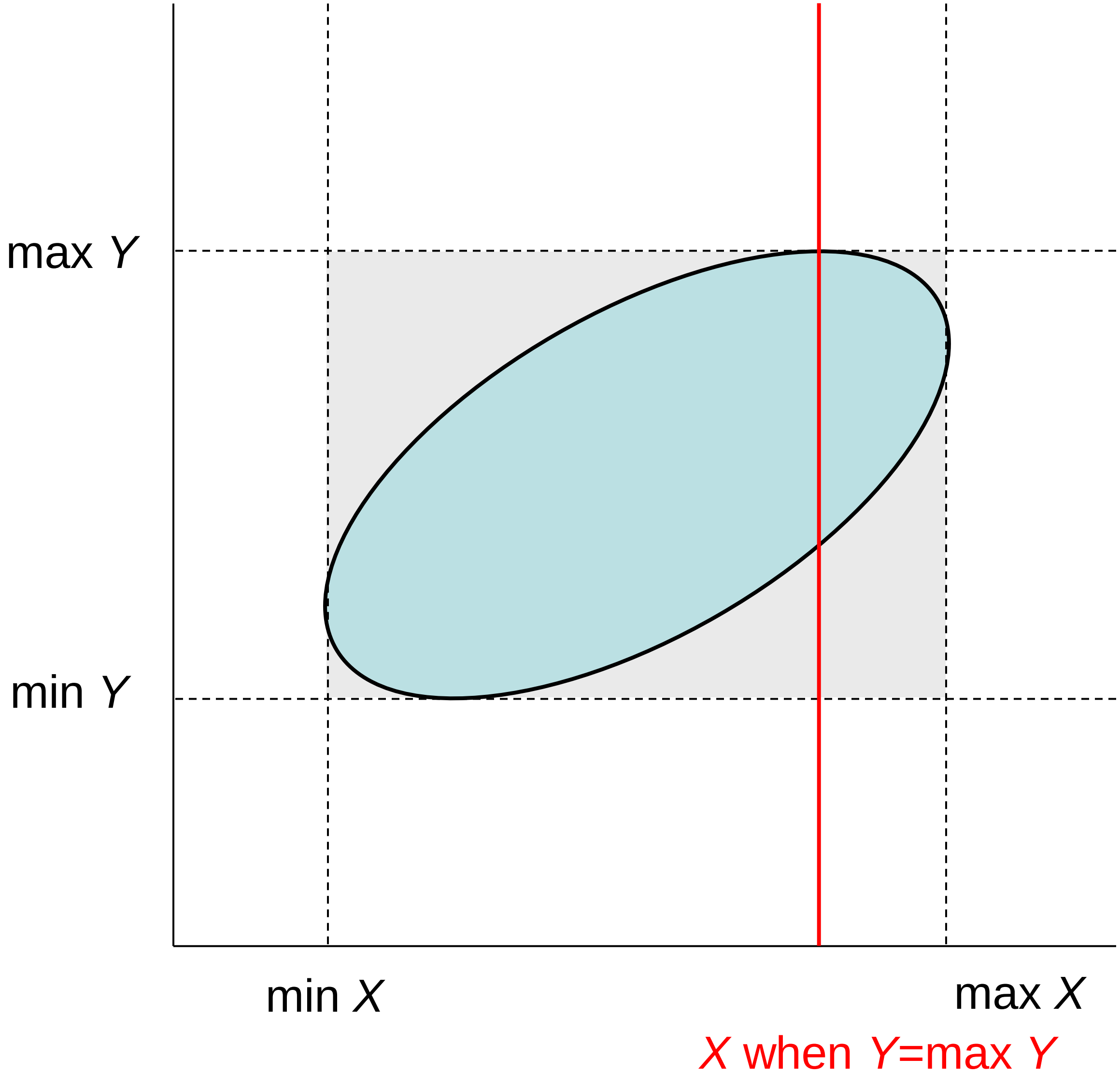
$$\frac{1}{\alpha_x^2}\left(x-\tilde{x}\right)^2-\frac{1}{\beta^2}\left(y-\tilde{y}\right)^2=1$$

$$\alpha=\frac{1-\beta\Delta_y^2}{2\Delta_x^2} \qquad \beta=\frac{1}{2\Delta_y^2}$$

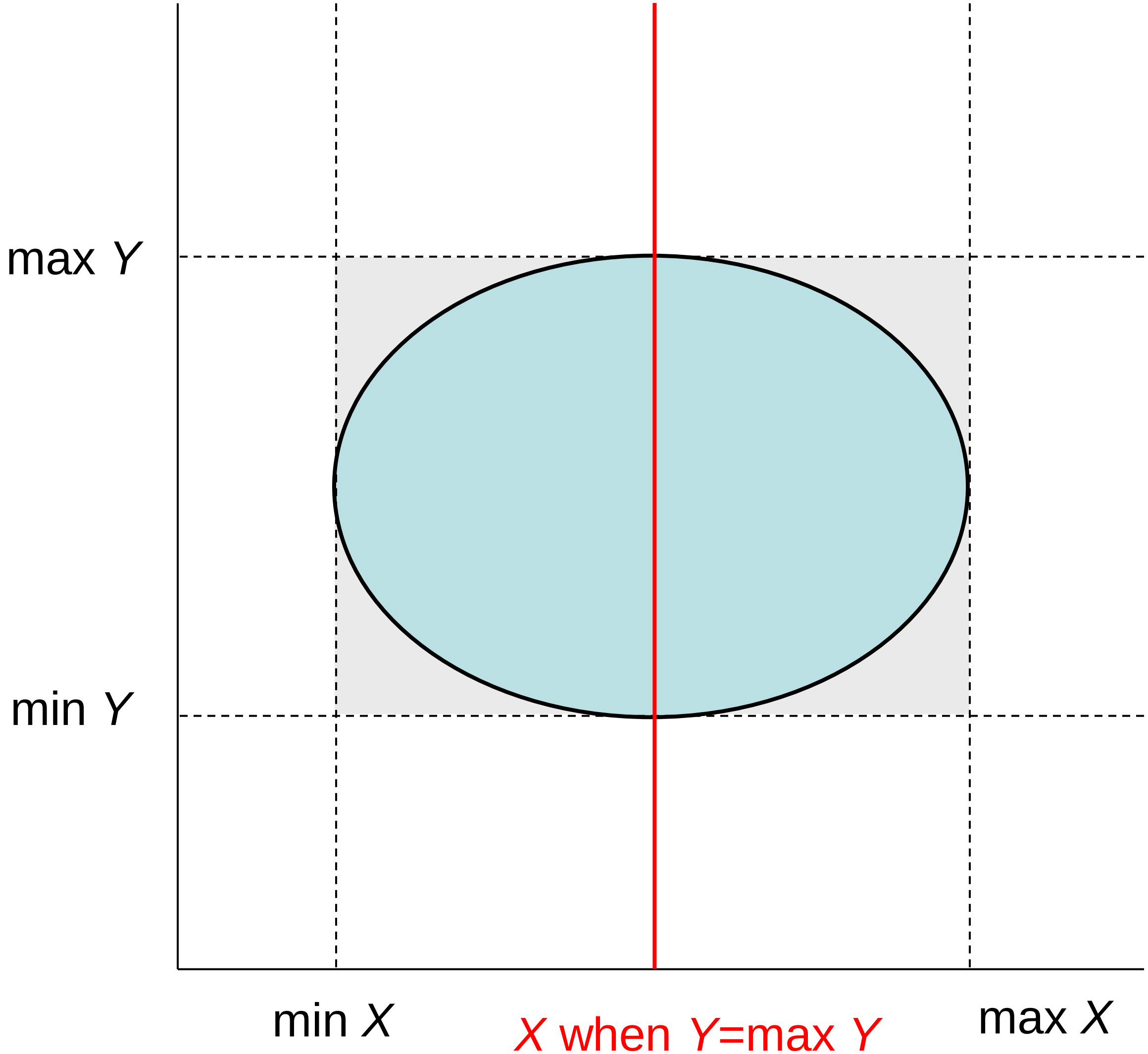
Strong correlation



Some correlation



No correlation



Fermi estimate of correlation

- Ask what value X takes when Y is at its max
- Pool values from multiple experts if needed
- The X value specifies the ellipse and thus the correlation between X and Y
- You can reverse the query: ask what value Y has when X is at its max; this gives a check

Fermi estimation

- Once marginal distributions and correlations are specified, Monte Carlo simulation can be used to compute the output distribution

- Results are contingent of several assumptions
  - Distribution shape assumptions
  - Correlation is a sufficient model of dependence
  - Approximate parameterizations are good enough

- Approach can't use other available information

Imprecise probabilities

Probability of an event

- Imagine a gamble that pays one dollar if an event occurs (but nothing otherwise)
  - How much would you pay to buy this gamble?
  - How much would you be willing to sell it for?
- Probability theory requires the same price for both
  - By asserting the probability of the event, you agree to buy any such gamble offered for this amount or less, *and* to sell the same gamble for any amount equal to or more than this 'fair' price…and to do so for *every* event!
- IP just says, sometimes, your highest buying price might be smaller than your lowest selling price

Why is yet another method needed?

- Some statements of uncertainty can’t be expressed with p-boxes or interval probability
- Need to express uncertainty of *all kinds*
- Sometimes mere bounding isn’t good enough

What bounding probability can't do

- Represent comparative probability judgments, e.g., event *A* is at least as likely as event *B*
- Give unique expectations needed for making decisions
- Give unique conditional probabilities needed for making inferences
- Maintain best possible bounds through updating

(Walley 2000)

Imprecise probabilities

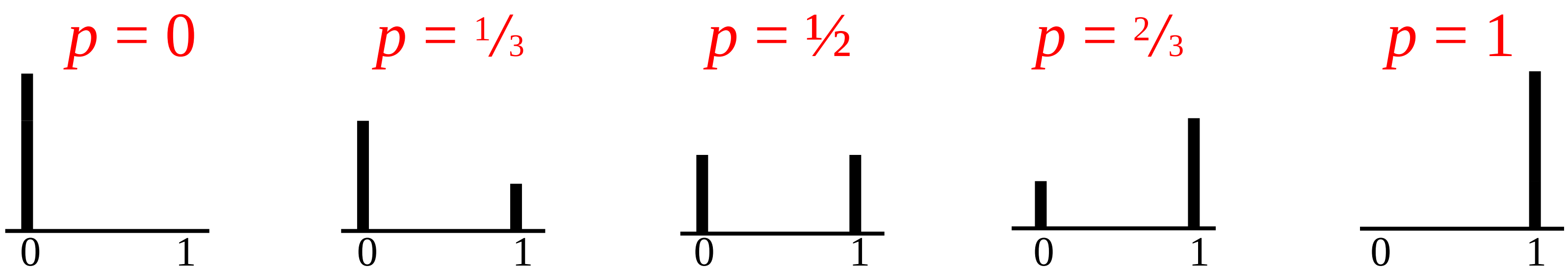
- Generic term for any theory that doesn’t assume a unique underlying probability
- Often expressed in terms of closed, convex sets of probability distributions (not the same as a p-box)
- General case v. special case

    set v. interval

    imprecise probabilities v. p-boxes, probability intervals

Sets of distribution functions

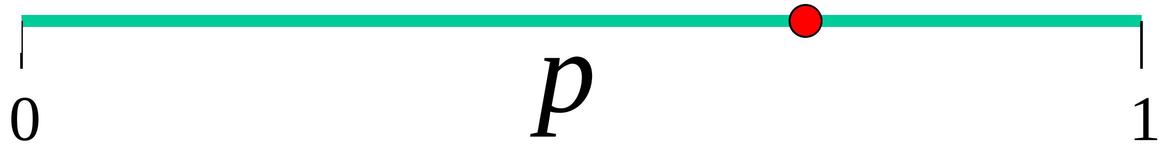
- Consider the set of all Bernoulli distributions (which are discrete with mass at only 0 and 1)



- Clearly, there's a one-dimensional family of such distributions, parameterized by how the mass is distributed between the two points

Space of distributions

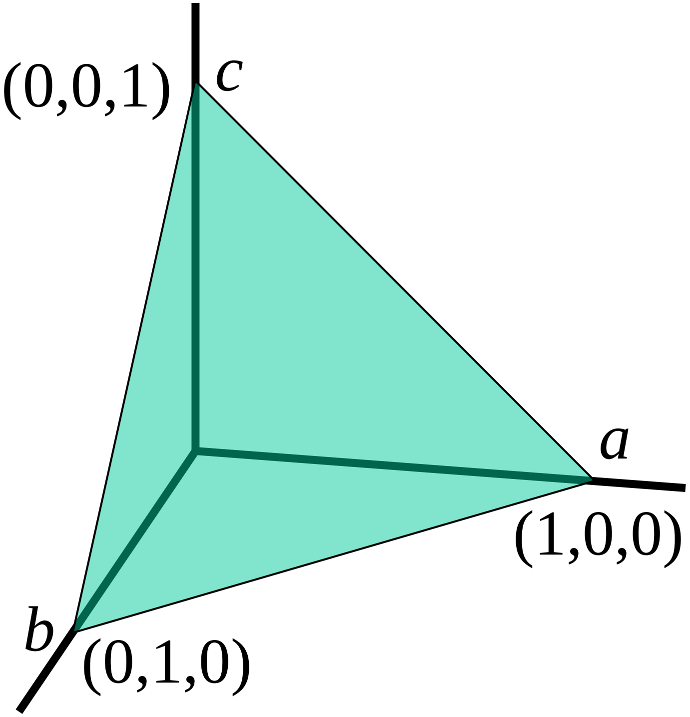
- This one-dimensional family constitutes a space of distributions in which each point represents a distribution



Three-dimensional case

- When the distributions have 3 point masses, the space becomes two-dimensional and has a triangular shape

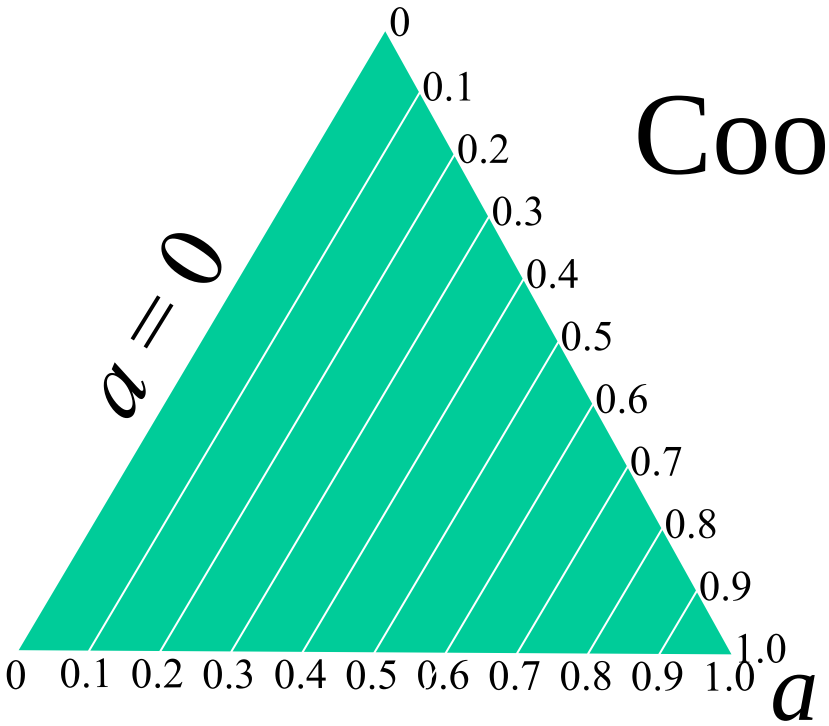
- The points on this surface are those whose coordinates add to one

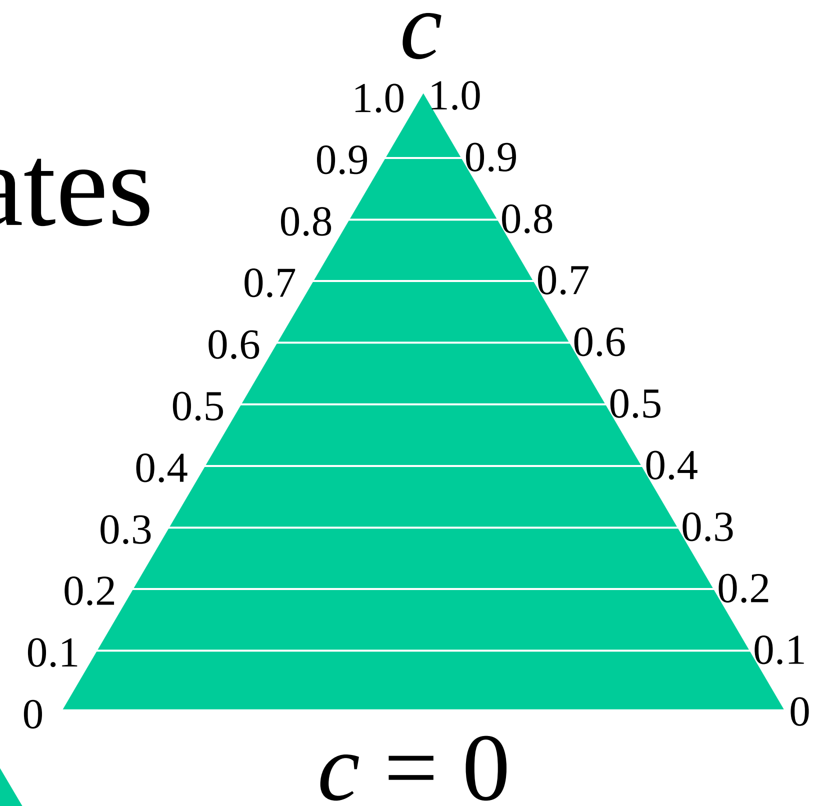


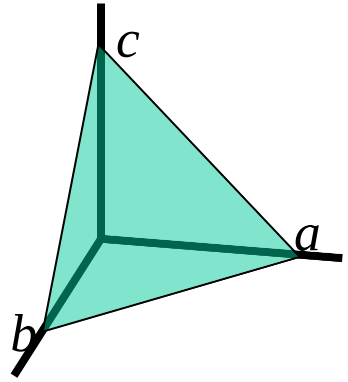
Simplex

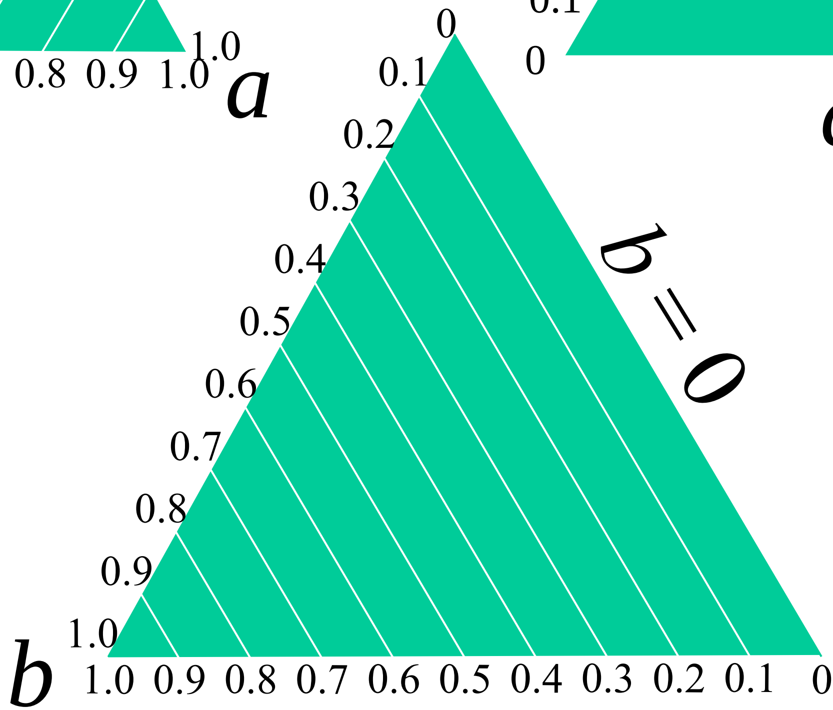
- For discrete distributions with n masses, the space, called a *simplex*, has $(n-1)$-dimensions
- One degree of freedom is lost to the constraint that probabilities sum to one
- For the continuous case, the space becomes infinite-dimensional, or you could be content to use discrete approximations

Coordinates









Walley's (2000) football game

- 3 possibilities for our team: win, draw, loss
- Suppose we have qualitative judgments:

  'Not win' is at least as probable as win

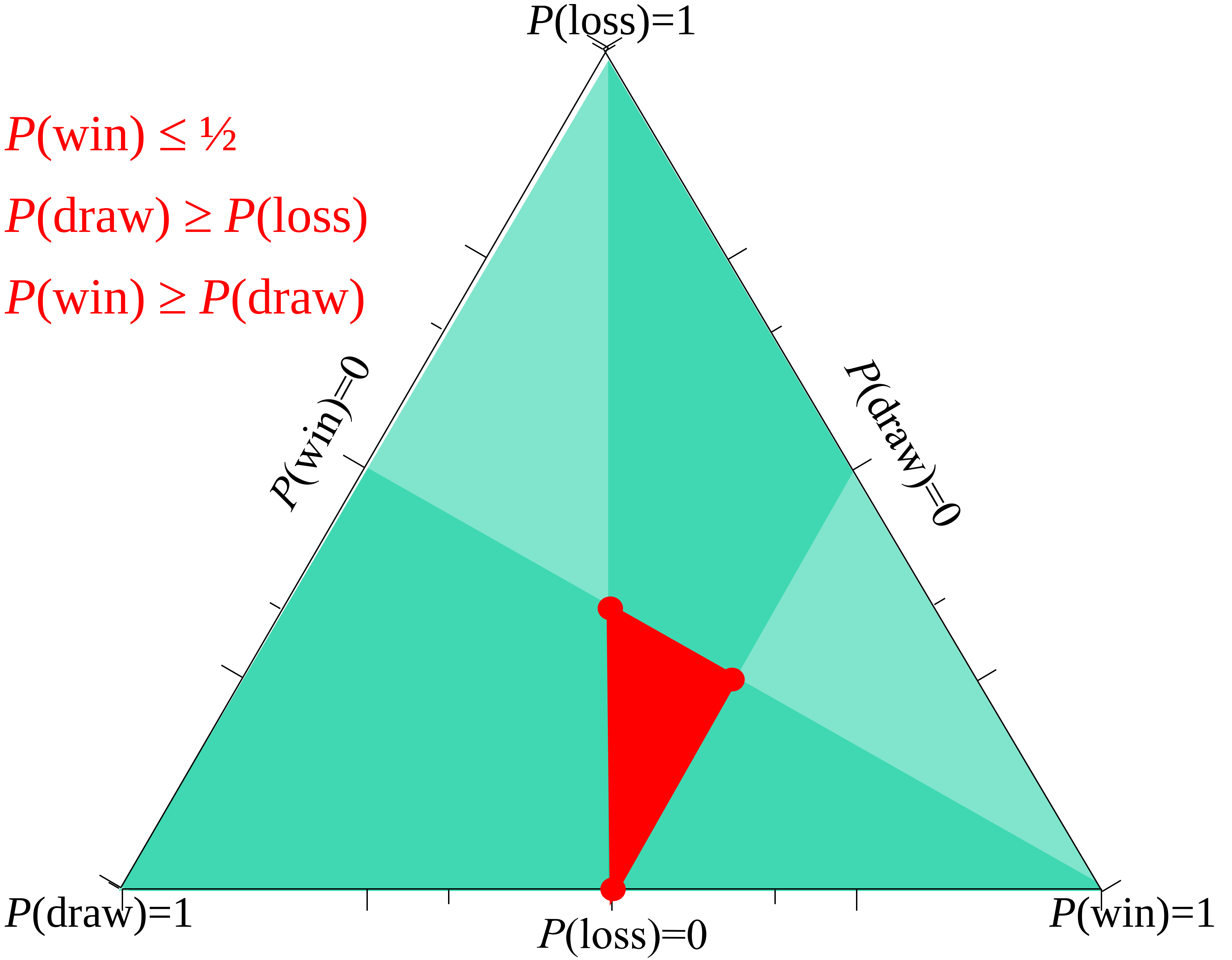
  Win is at least as probable as draw

  Draw is at least as probable as loss
- These constrain the probability distribution P

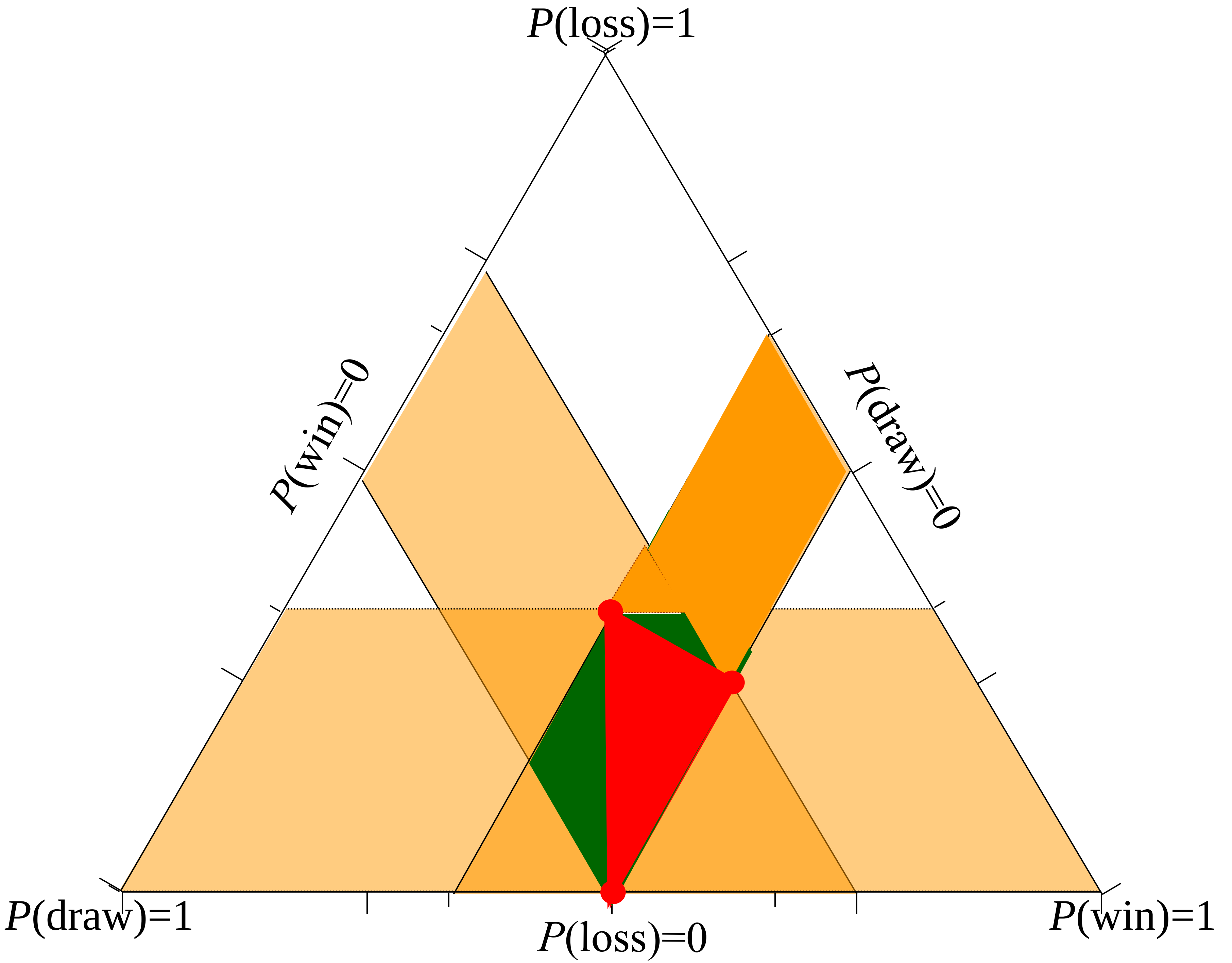
  $P(\text{win}) \leq \frac{1}{2}$

  $P(\text{win}) \geq P(\text{draw})$

  $P(\text{draw}) \geq P(\text{loss})$



P(win) ≤ ½
P(draw) ≥ P(loss)
P(win) ≥ P(draw)
P(loss)=1
P(win)=0
P(draw)=0
P(draw)=1
P(loss)=0
P(win)=1



P(loss)=1
P(win)=0
P(draw)=0
P(draw)=1
P(loss)=0
P(win)=1

So what?

- The closed, convex set of probability distributions (the red triangular region) expresses the uncertainty
- This set of distributions is smaller than the set implied by bounds on the three probabilities (the green area enclosing the triangle)
- This difference can affect expectations of functions that depend on the events, and conditional probabilities

Credal set

- Knowledge and judgments can be used to define a set of possible probability measures

Kinds of information

  - All distributions within bounds are possible
  - Only distributions having a given shape
  - Probability of an event is within some interval
  - Event *A* is at least as probable as event *B*
  - Nothing is known about the probability of *C*

- Computing with credal sets usually requires mathematical programming

Case studies

Case study:

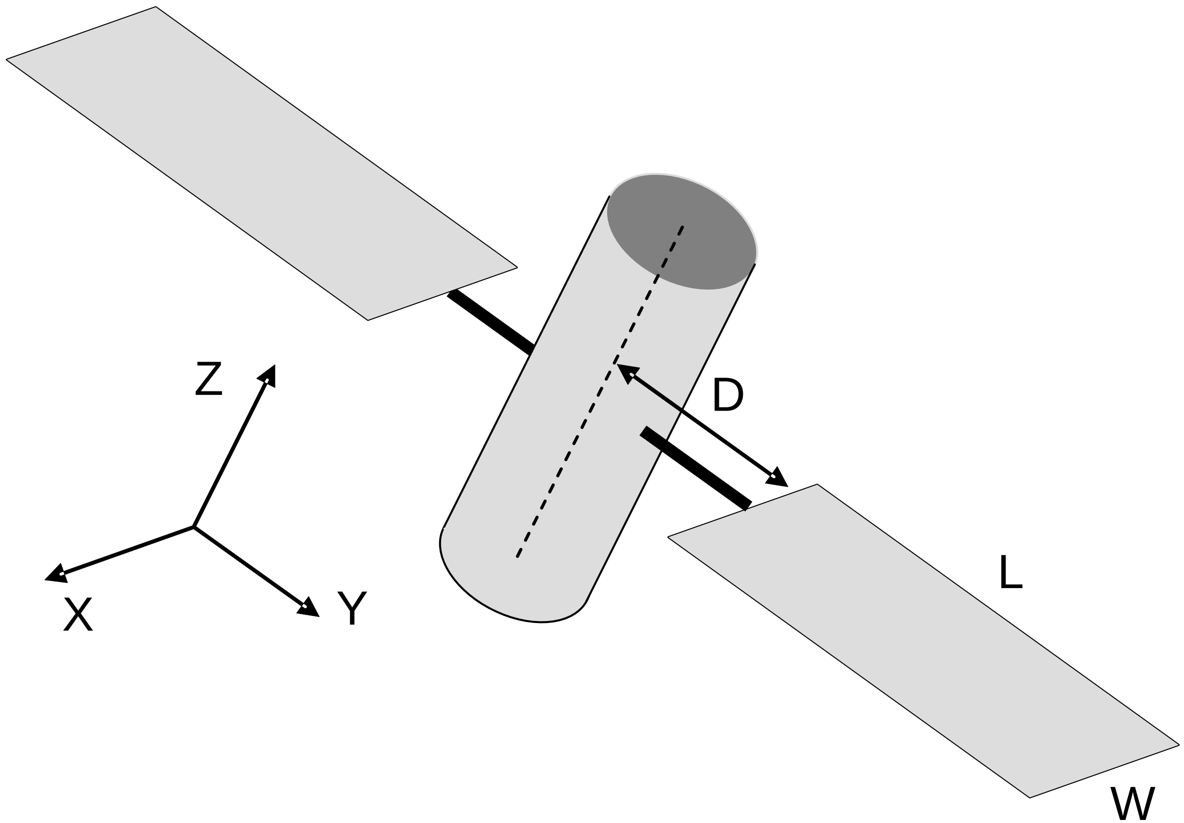
Spacecraft design under mission uncertainty

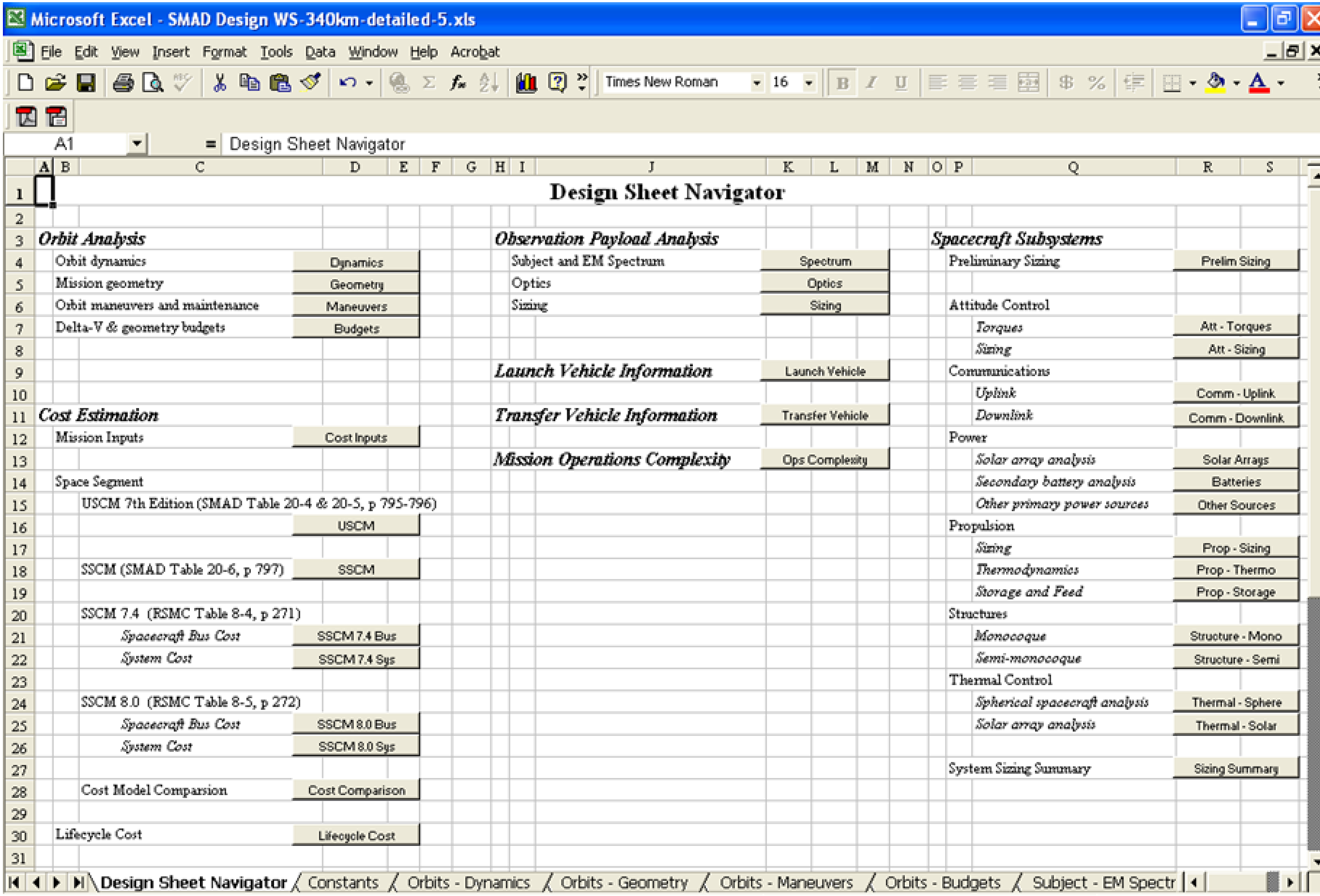
Integrated concurrent engineering

- Real-time collaborative, interactive process
  - Implemented at JPL, LaRC, and others
- Reduces design time by an order of magnitude
  - But quantitative risk assessment is difficult
- Design solutions are iterative
  - So a Monte Carlo approach may not be practical

Mission

Deploy satellite carrying a large optical sensor





Wertz and Larson (1999) *Space Mission Analysis and Design* (***SMAD***). Kluwer.

Microsoft Excel - SMAD Design WS-340km-detailed-5.xls

File Edit View Insert Format Tools Data Window Help Acrobat

M17 =

Return to Navigator

Attitude Control - Torque Estimates

(All information on this sheet is contained in the block from Cell A1 to Cell Q27)

Orbit characteristics			
Altitude		**340.000**	**km**
Satellite velocity		**7.703**	**km/s**
Atmospheric density		**1.983E-11**	**kg/m^3**
Sun incidence angle		**0.00**	**deg**
Maximum deviation from local vertical		**10.00**	**deg**
Spacecraft characteristics			
Largest moment of inertia	*7315.000*	**7315.000**	**kg-m^2**
Smallest moment of inertia	*4655.000*	**4655.000**	**kg-m^2**
Projected surface area	*14.063*	**14.063**	**m^2**
Moment arm for solar radiation torques		**0.250**	**m**
Moment arm for aerodynamic torques		**0.250**	**m**
Drag coefficient		**3.13**	
Surface reflectivity		**0.60**	
Residual dipole		**1.00**	**A-m^2**
Slew characteristics			
Maximum slewing angle	*38.00*	**38.00**	**deg**
Minimum maneuver time	*760.00*	**760.00**	**sec**

Environmental torques		
Gravity gradient	**1.794E-03**	**N-m**
Solar radiation	**2.565E-05**	**N-m**
Magnetic	**5.250E-05**	**N-m**
Aerodynamic	**6.473E-03**	**N-m**
Total (RSS)	**6.717E-03**	**N-m**
Slewing torque	**3.360E-02**	**N-m**

Obs Payload - Optics / Obs Payload - Sizing / SC Bus - Preliminary Sizing / **Attitude Control - Torques** / Attitude Control - Sizing / Comm

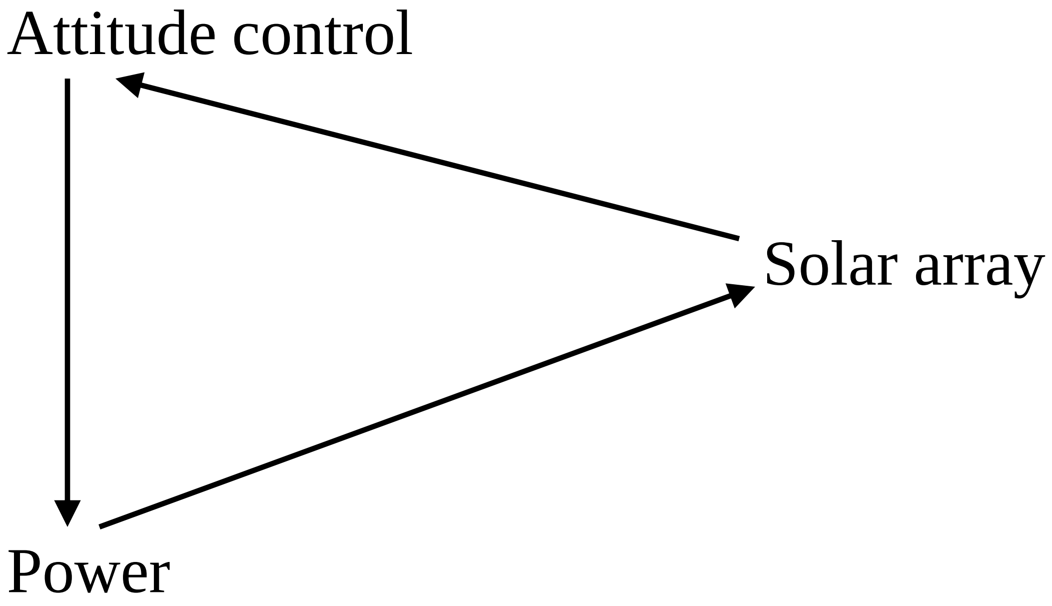
Typical subsystems

Attitude control
Command data systems
Configuration
Cost
Ground systems
Instruments
Mission design
Power

Program management
Propulsion
Science
Solar array
Systems engineering
Telecommunications – System
Telecommunications – Hardware
Thermal control

Moments of inertia determine
Torque needed determines
Power requirements determines
Solar panel size determines
Moments of inertia

Demonstration system



- Calculations within a single subsystem (ACS)
- Calculations within linked subsystems

Attitude control subsystem (ACS)

- 3 reaction wheels
- Design problem: solve for *h*
  - Required angular momentum
  - Needed to choose reaction wheels
- Mission constraints
  - Δt_{orbit} = 1/4 orbit time
  - θ_{slew} = max slew angle
  - Δt_{slew} = min maneuver time
- Inputs from other subsystems
  - I, I_{max}, I_{min} = inertial moment
  - Depend on solar panel size, which depends on power needed, so on *h*

$$h = \tau_{tot} \times \Delta t_{orbit}$$

$$\tau_{tot} = \tau_{slew} + \tau_{dist}$$

$$\tau_{slew} = \frac{4\theta_{slew}}{\Delta t_{slew}^2} I$$

$$\tau_{dist} = \tau_g + \tau_{sp} + \tau_m + \tau_a$$

$$\tau_g = \frac{3\mu}{2(R_E + H)^3} \left| I_{max} - I_{min} \right| \sin(2\theta)$$

$$\tau_{sp} = L_{sp} \frac{F_s}{c} A_s (1 + q) \cos(i)$$

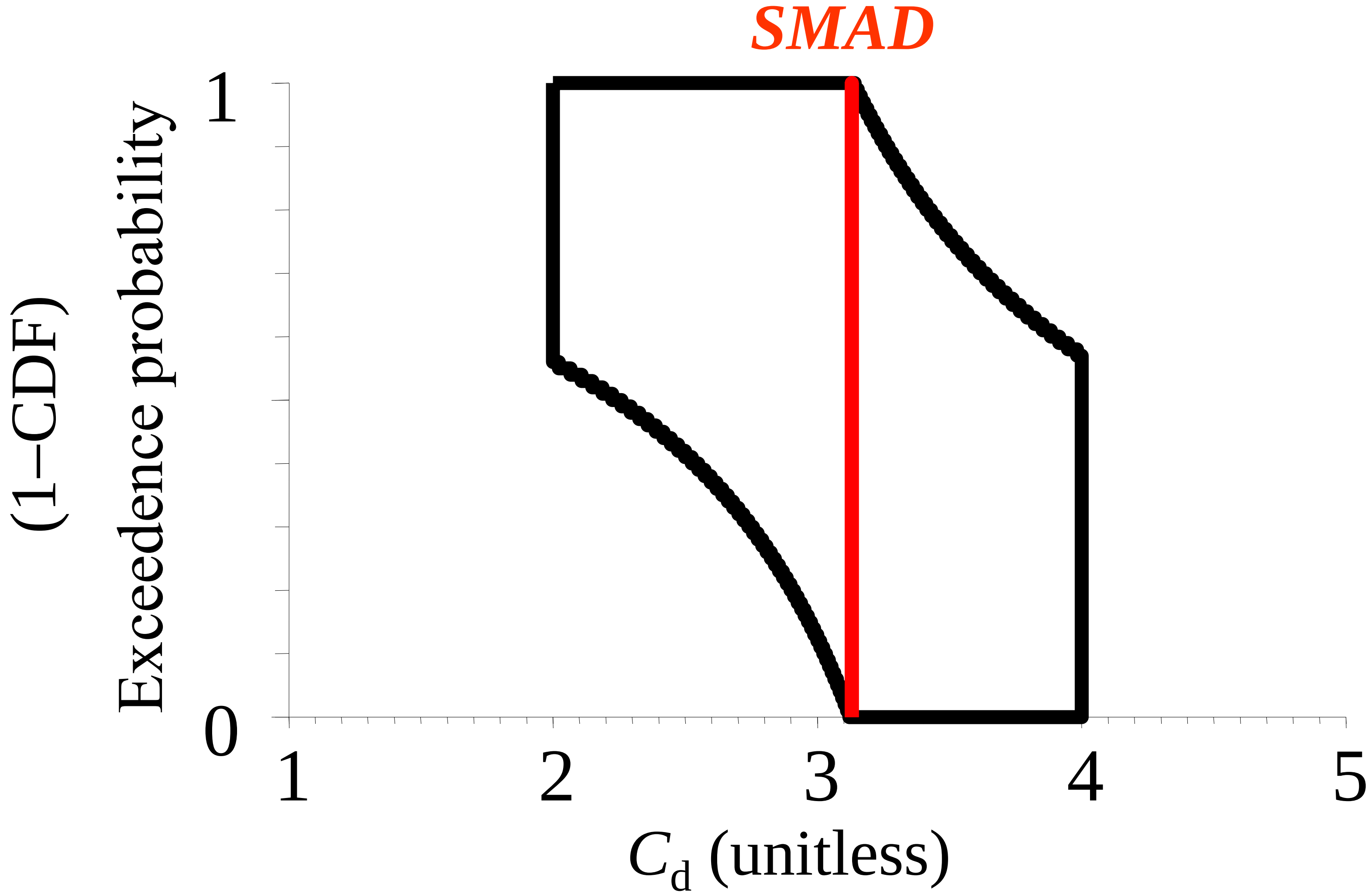
$$\tau_m = \frac{2MD}{(R_E + H)^3}$$

$$\tau_a = \frac{1}{2} L_a \rho C_d A V^2$$

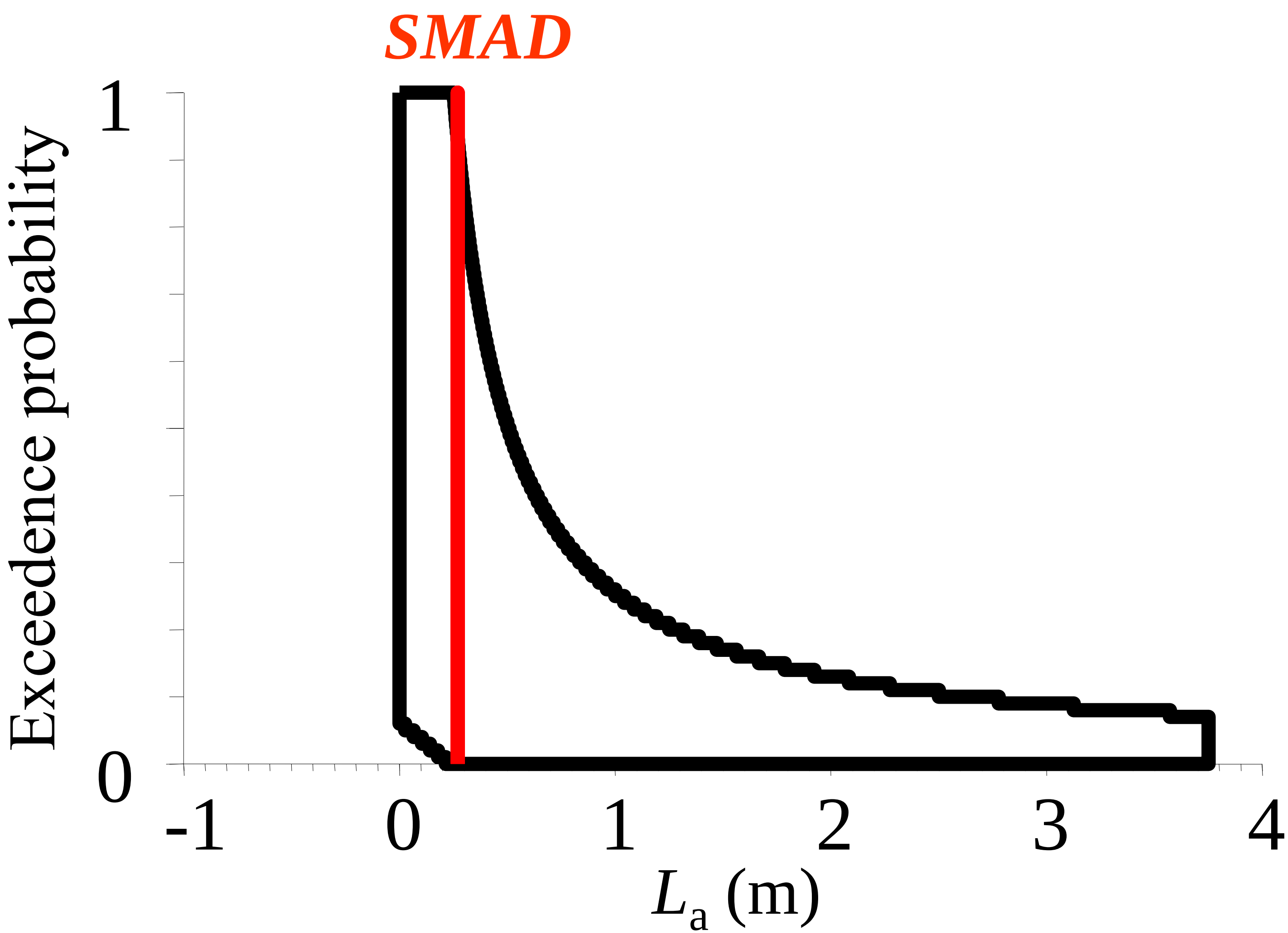
Attitude control input variables

Symbol	**Unit**	**Variable**	**Type**	**Value**	***SMAD***
C_d	unitless	Drag coefficient	p-box	range=[2,4] mean=3.13	3.13
L_a	m	Aerodynamic drag torque moment	p-box	range=[0,3.75] mean=0.25	0.25
L_{sp}	m	Solar radiation torque moment	p-box	range=[0,3.75] mean=[0.25]	0.25
D	A m^2	Residual dipole	interval	[0,1]	1
i	degrees	Sun incidence angle	interval	[0,90]	0
ρ	kg m^3	Atmospheric density	interval	[3.96e-12, 9.9e-11]	1.98e-11
θ	degrees	Major moment axis deviation from nadir	interval	[10,19]	10
q	unitless	Surface reflectivity	interval	[0.1,0.99]	0.6
I_{min}	kg m^2	Minimum moment of inertia	interval	[4655]	4655
I_{max}	kg m^2	Maximum moment of inertia	interval	[7315]	7315
μ	m^3 s^{-2}	Earth gravity constant	point	3.98e14	3.98e14
A	m^2	Area in the direction of flight	point	3.75^2	3.75^2
RE	km	Earth radius	point	6378.14	6378.14
H	km	Orbit altitude	point	340	340
F_s	W m^{-2}	Average solar flux	point	1367	1367
θ_{slew}	degrees	Maximum slewing angle	point	38	38
c	m s^{-1}	Light speed	point	2.9979e8	2.9979e8
M	A m^2	Earth magnetic moment	point	7.96e22	7.96e22
Δt_{slew}	s	Minimum maneuver time	point	760	760
A_s	m^2	Area reflecting solar radiation	point	3.75^2	3.75^2
Δt_{orbit}	s	Quarter orbit period	point	1370	1370

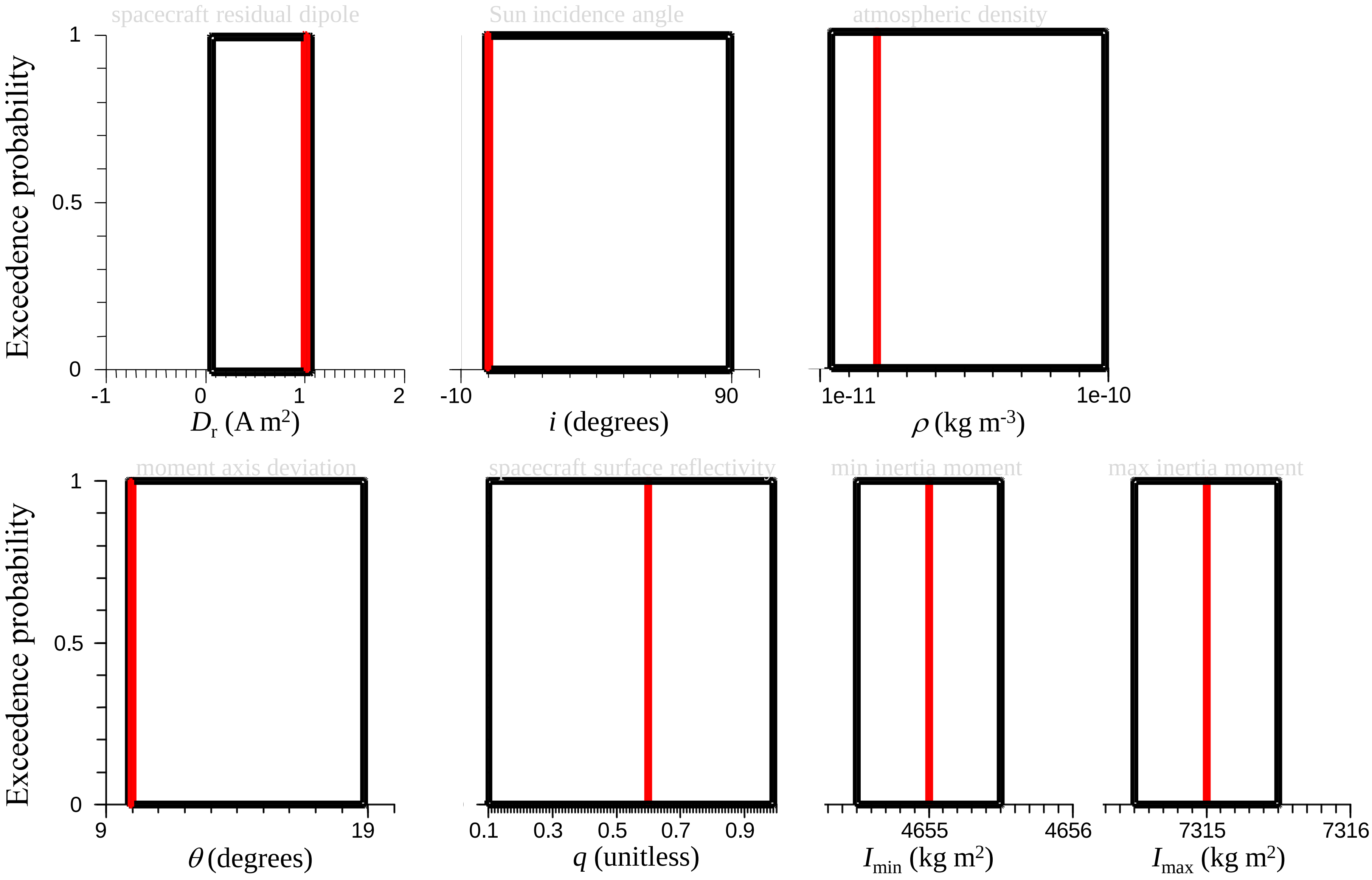
Coefficient of drag, C_d



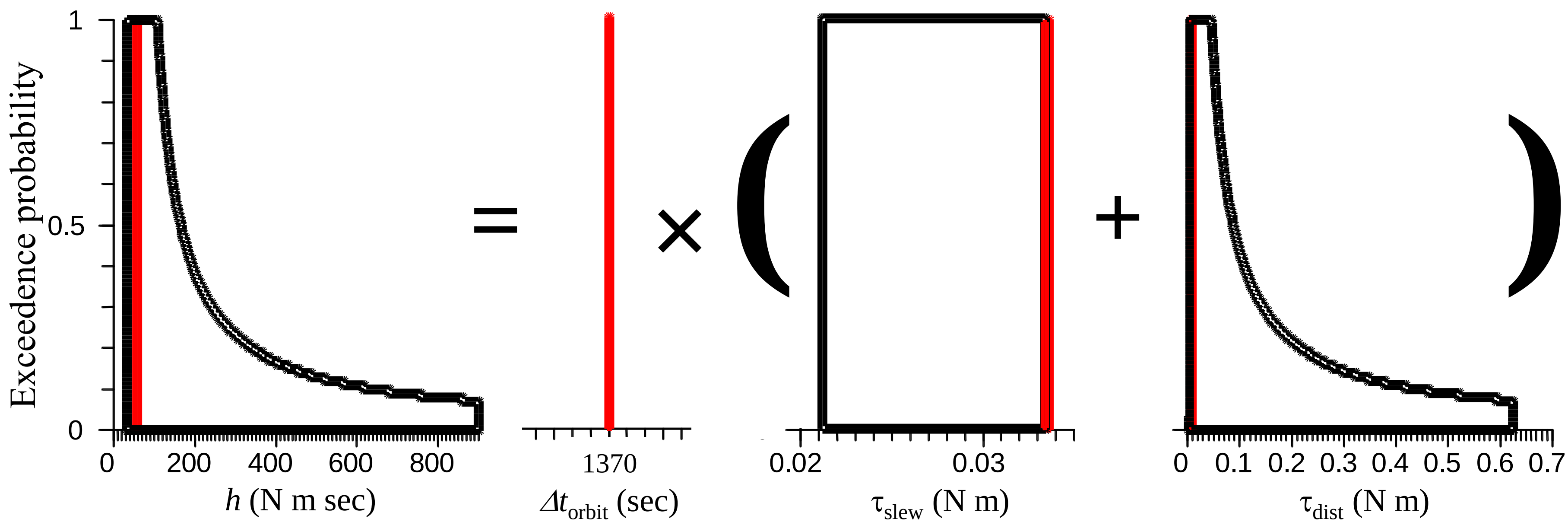
Aerodynamic drag torque moment, L_a



Interval inputs and *SMAD* points

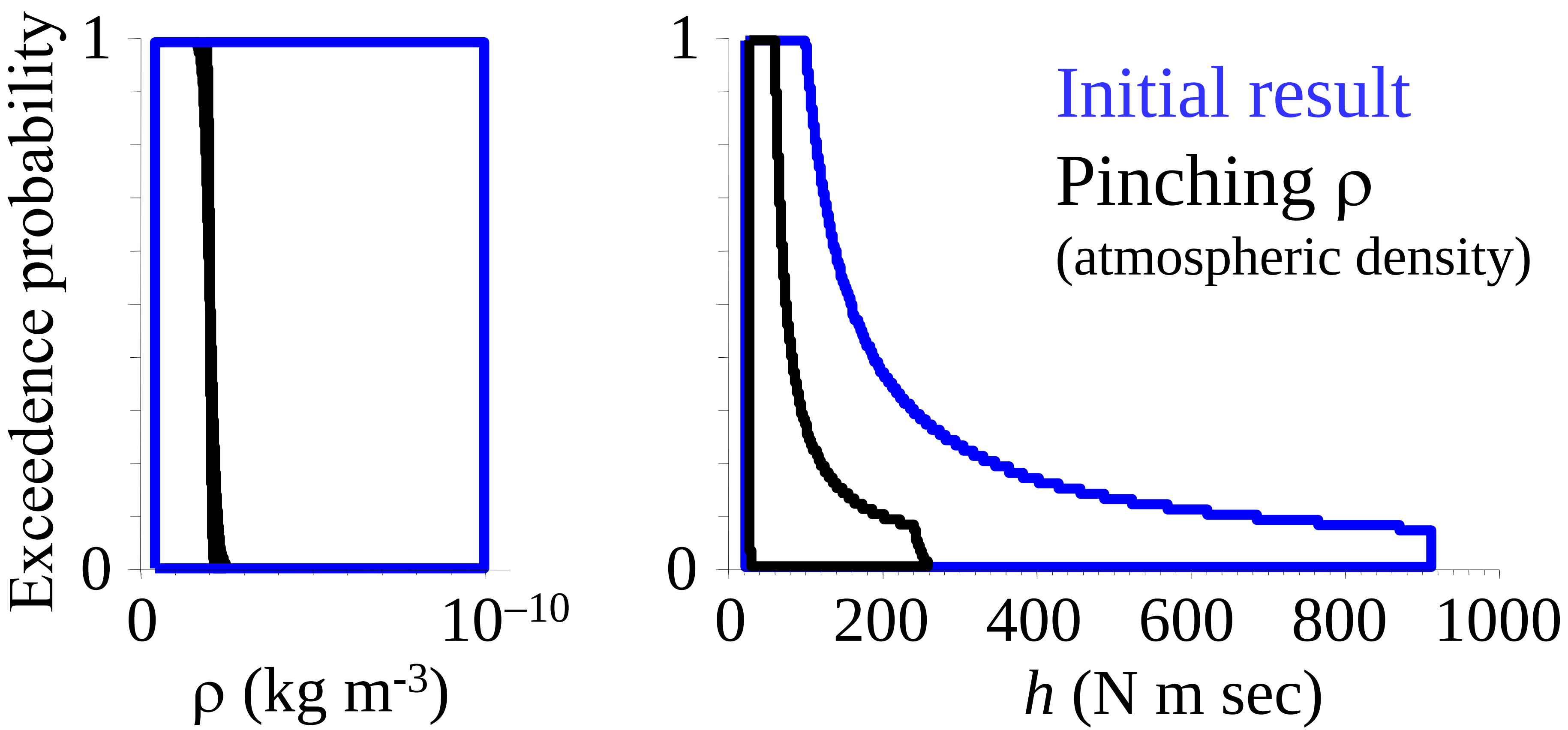


Required angular momentum, *h*



$$h = \Delta t_{orbit} \times \left(\tau_{slew} + \tau_{dist} \right)$$

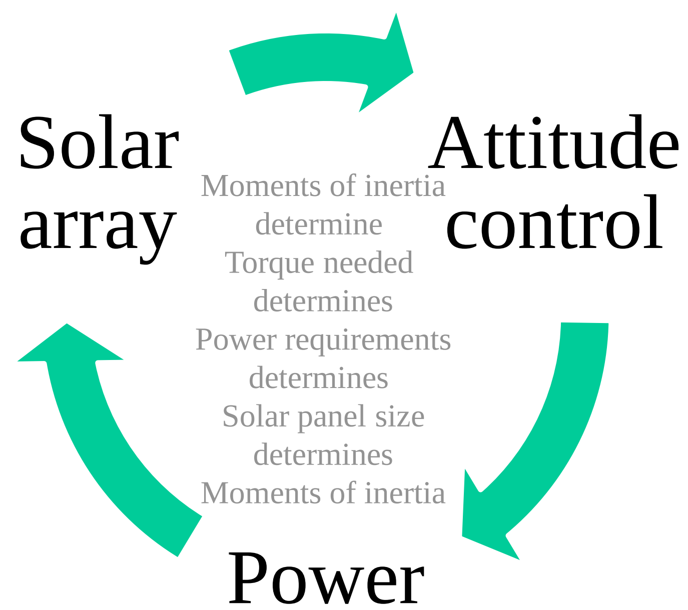
Value of information: pinching ρ



Linked subsystems

- Minimum moment of inertia I_{min}
- Maximum moment of inertia I_{max}
- Total torque τ_{tot}
- Total power P_{tot}
- Solar panel area A_{sa}

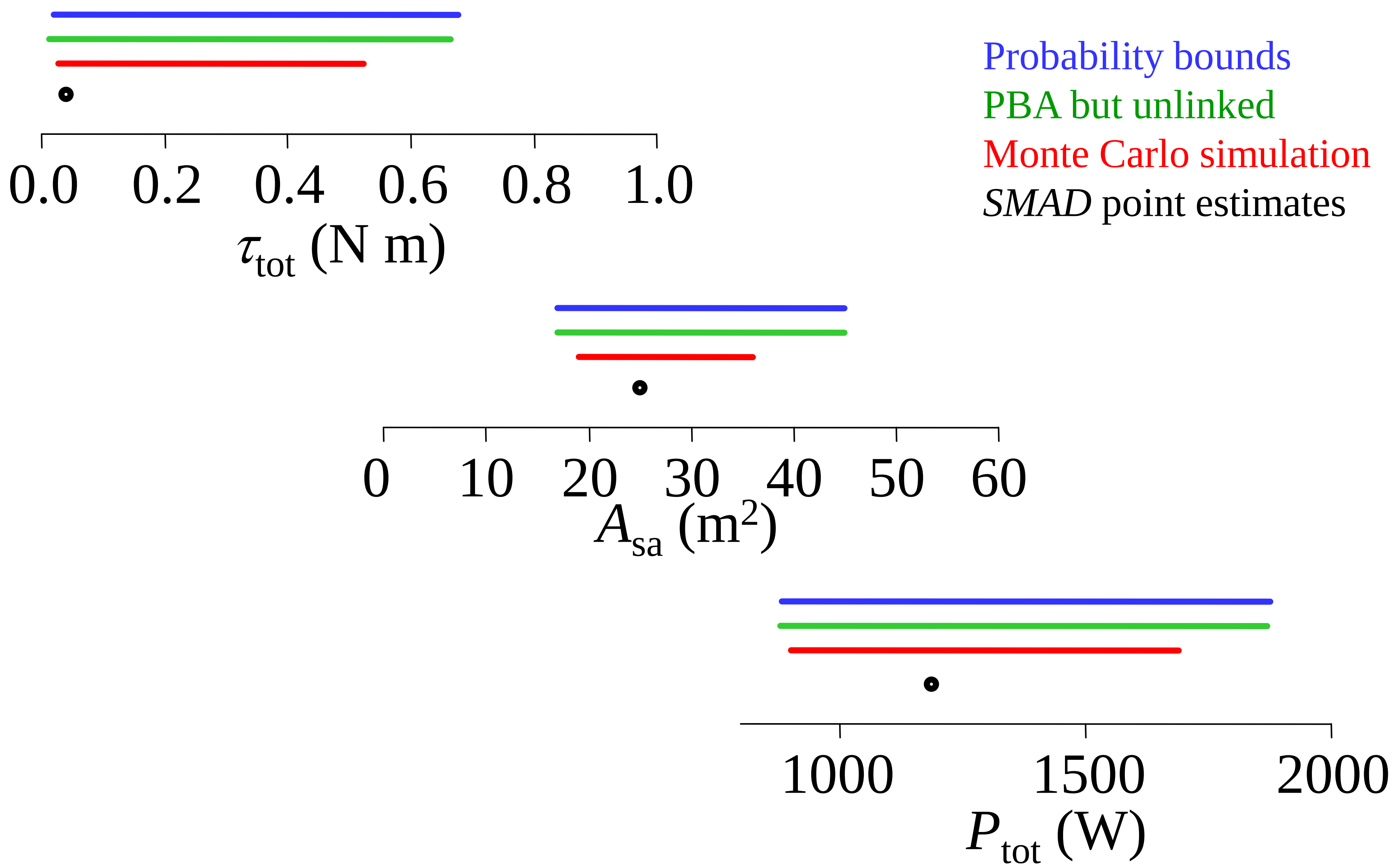
Iteratively calculated



Analysis of calculations

- Do p-boxes enclose Monte Carlo and *SMAD*?
- Does iteration through links cause runaway uncertainty growth (or contraction)?
- Four parallel analyses
  - *SMAD*'s point estimates
  - Monte Carlo simulation
  - P-boxes but without linkage among subsystems
  - P-boxes with fully linked subsystems

Ranges of results



Findings

- Calculations workable
  - No runaway inflation (or loss) of uncertainty
  - Comprehensive bounds easier than via Monte Carlo
- Practical and useful results
  - Uncertainty influences engineering decisions
  - Reducing uncertainty about ρ (by picking a launch date) strongly reduces design uncertainty

Attitude control subsystem

Symbol	Unit	Type	Value	SMAD
C_d	unitless	p-box	range=[2,4] mean=3.13	3.13
L_a	m	p-box	range=[0,3.75] mean=0.25	0.25
L_{sp}	m	p-box	range=[0,3.75] mean=[0.25]	0.25
D	A m^2	interval	[0,1]	1
i	degrees	interval	[0,90]	0
ρ	kg m^3	interval	[3.96e-12, 9.9e-11]	1.98e-11
θ	degrees	interval	[10,19]	10
q	unitless	interval	[0.1,0.99]	0.6
I_{min}	kg m^2	interval	[4655]	4655
I_{max}	kg m^2	interval	[7315]	7315
μ	m^3 s^{-2}	point	3.98e14	3.98e14
A	m^2	point	3.75^2	3.75^2
RE	km	point	6378.14	6378.14
H	km	point	340	340
F_s	W m^{-2}	point	1367	1367
θ_{slew}	degrees	point	38	38
c	m s^{-1}	point	2.9979e8	2.9979e8
M	A m^2	point	7.96e22	7.96e22
Δt_{slew}	s	point	760	760
A_s	m^2	point	3.75^2	3.75^2
Δt_{orbit}	s	point	1370	1370

$$h = \tau_{\mathrm{tot}} \times \Delta t_{\mathrm{orbit}}$$

$$\tau_{tot} = \tau_{slew} + \tau_{dist}$$

$$\tau_{\mathrm{slew}} = \frac{4\theta_{\mathrm{slew}}}{\Delta t_{\mathrm{slew}}^{\ 2}} I$$

$$\tau_{dist} = \tau_g + \tau_{sp} + \tau_m + \tau_a$$

$$\tau_{\mathrm{g}} = \frac{3\mu}{2(R_{\mathrm{E}} + H)^3} |I_{\mathrm{max}} - I_{\mathrm{min}}| \sin(2\theta)$$

$$\tau_{\mathrm{sp}} = L_{\mathrm{sp}} \frac{F_{\mathrm{s}}}{c} A_{\mathrm{s}} (1+q) \cos(i)$$

$$\tau_{\mathrm{m}} = \frac{2MD}{(R_{\mathrm{E}} + H)^3}$$

$$\tau_{\mathrm{a}} = \frac{1}{2} L_{\mathrm{a}} \rho C_{\mathrm{d}} A V^2$$

Be careful with units!

Case study:

Risks from eating fish caught in the Housatonic River

Abstract

We conducted point and probabilistic risk assessments for the consumption of fish in the Housatonic River. The probabilistic analyses consisted of Monte Carlo simulation (with both one-dimensional and microexposure event analyses), dependency bounds analyses, probability bounds analyses, and a comparison of the results of a probability bounds analysis with a 2 dimensional Monte Carlo approach. The exposure distributions were based on the same datasets from which the point estimate exposure parameters were derived. The Monte Carlo analyses assumed strict independence between all variables and indicated the expected variability of the risk. A dependency bounds analysis was used to relax the assumptions of independence and explore risks from ingestion under other dependency assumptions. The probability bounds analyses used intervals to comprehensively bound the uncertainty in the distribution of the risk indicated by the Monte Carlo simulations. Sensitivity analyses were conducted to quantify the contributions of variability and/or uncertainty of each exposure parameter to the risk estimates. The ingestion rate usually contributed the largest amount of variance in the models examined.

Outline of analyses

Probabilistic analyses

- One-dimensional (1D) and Microexposure model
- Cancer risk and non-cancer risk
- Monte Carlo simulation (MC)

  variability only
- Probability bounds analysis (PBA)

  variability, uncertainty
- Dependency bounds analysis (DBA)

Cancer models

1-D model:

$$Dose = \frac{C_{fish} \times (1 - LOSS) \times FI \times CF \times \left(\frac{ED_C \times EFIR_C}{BW_C} + \frac{ED_A \times EFIR_A}{BW_A} \right)}{AT}$$

Microexposure model:

$$Dose = \frac{CF}{AT} \left[\left(\frac{ED_C}{BW_C} \times mean(FI \times EF) \times Z_C \right) + \left(\frac{ED_A}{BW_A} \times mean(FI \times EF) \times Z_A \right) \right]$$

where $Z_j = mean\left(C_{fish} \times (1 - LOSS) \times IR_j\right)$ for j = C (child) or A (adult)

Input variables

Variable	Symbol	Units	Min,Max	Central Estimate	Std. Dev	MC Distribution	P-box
Concentration	C_{fish}	mg/kg	10.8, 13.9	13.9	-	point est.	interval
Cooking loss	*LOSS*	unitless	0, 1	0.26	0.18	mixture	mixture
Bake			0.05, 0.67	0.22	0.112	lognormal	MMMS
Broil			0.02, 1	0.19	0.18	T-lognormal	MMMS
Pan fry			0.04, 0.9	0.24	0.15	lognormal	MMMS
Deep fat fry			0.15, 1	0.44	0.17	T-lognormal	MMMS
Adult intake rate	$EFIR_A$	g/day	0.03, 647	8.5	13.6	EDF	ENV EDF
Child intake rate	$EFIR_C$	g/day	0.015, 324	4.25	6.8	EDF	ENV EDF
Adult ingestion rate	IR_A	g/meal	142, 340	227	-	triangular	interval
Child ingestion rate	IR_C	g/meal	70.9, 170	113.5	-	triangular	interval
Exposure frequency	*EF*	meal/yr	0.03, 490	13.1	22.2	decon. EDF	ENV decon. EDF
Fraction ingested	*FI*	unitless	0.1, 1	0.48	0.27	EDF	MMMS
Adult exposure dur.	ED_A	yr	1, 64	29	20	T-lognormal	MMMS
Child exposure dur.	ED_C	yr	1, 6	3.5	1.4	uniform	interval
Adult body weight	BW_A	kg	39, 119	72	15	lognormal	lognormal
Child body weight	BW_C	kg	12, 23	17	2.3	lognormal	lognormal

Variables indicated in red are needed for MEE modeling and are not discussed in this presentation.

Monte Carlo methods

- Precise point estimate
- Empirical distribution function (EDF)
- Precise distribution function
  - Lognormal
  - Truncated lognormal
  - Uniform
  - Triangular (MEE models only)
- Stochastic mixture of distributions

Probability bounding methods

- Interval (Min, Max)
- MMMS (Min, Max, Mean, Std.Dev)
- Envelope of EDFs
- Stochastic mixture of MMMS
- Precise distribution function (lognormal)

Concentration in fish (C_{fish})

- MC used the EPC for 4 fish species (combined)

  95% UCL used to "account for uncertainty"

  Using the UCL is intended to prevent underestimation

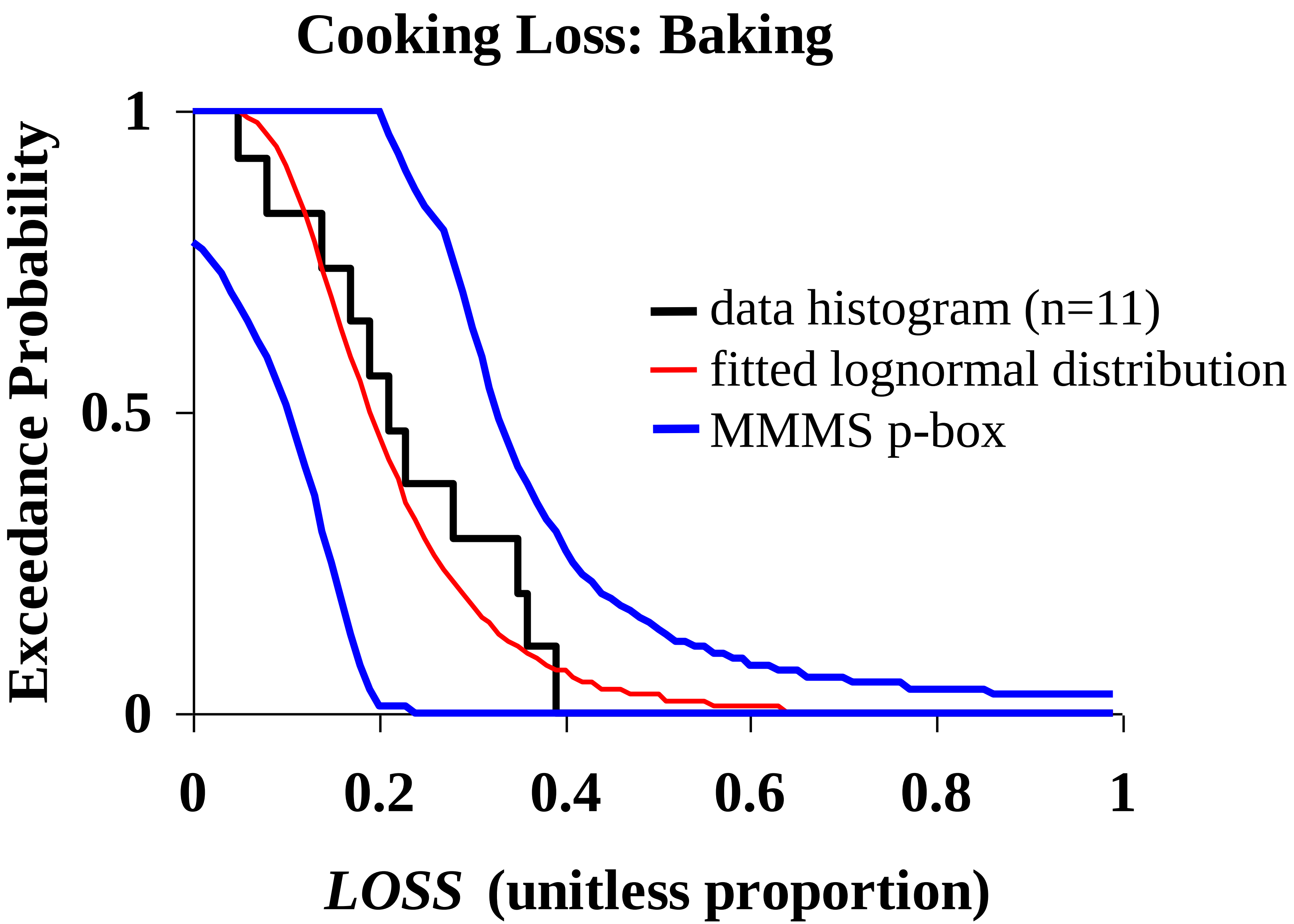
- PBA used the interval: [mean, EPC]

  Assumes the sample mean ≤ the true mean

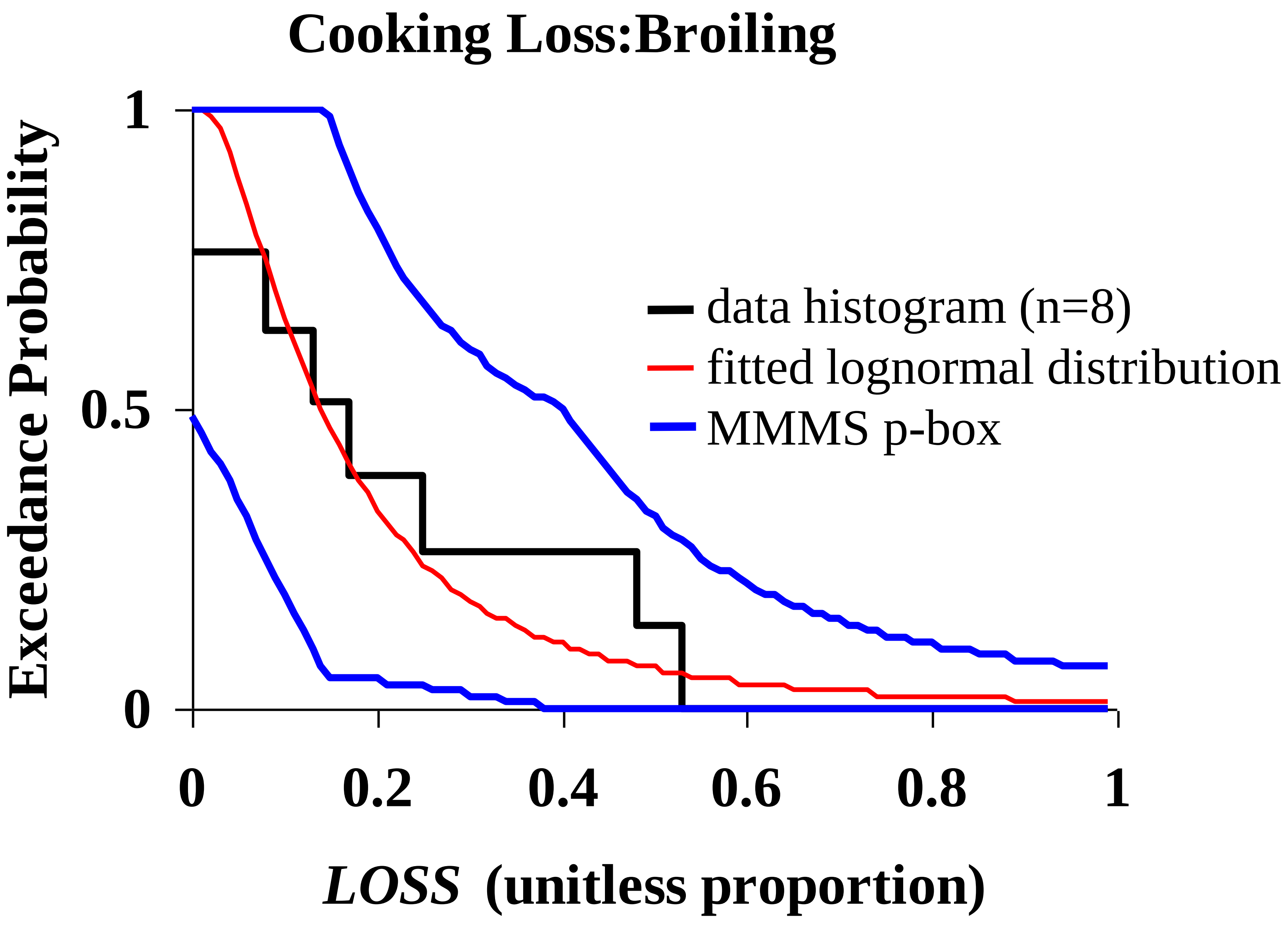
  EPA is concerned with not underestimating the mean

Cooking loss data

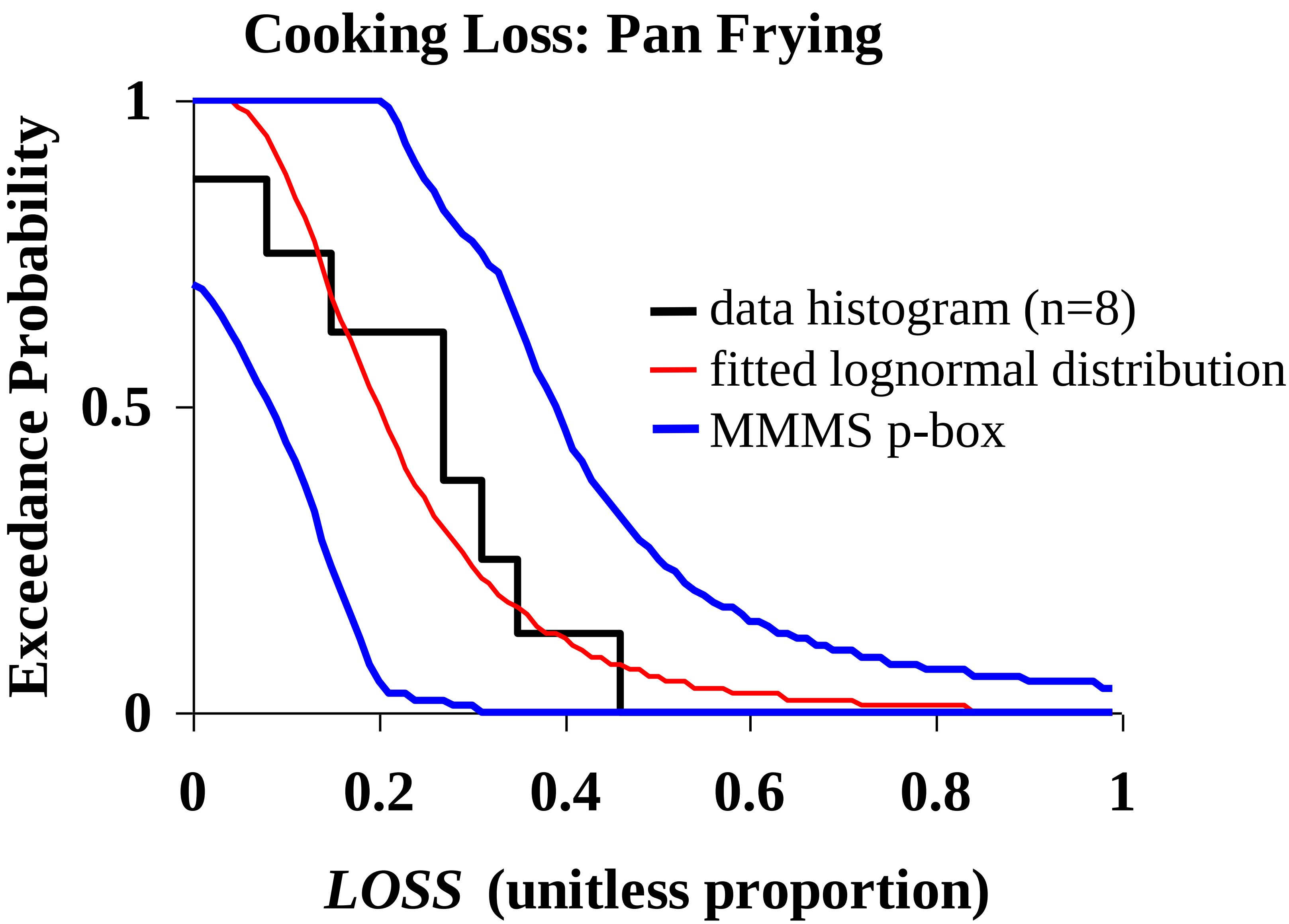
	Bake	Ref	Broil	Ref	Pan fry	Ref	Deep fat fry	Ref
	5	(8)	0	(7)	46	(4)	74	(7)
	16	(7)	53	(11)	7.5	(2)	31	(12)
	34	(11)	7.5	(2)	35	(13)	35	(13)
	7.5	(2)	24	(15)	31	(15)	32	(14)
	27	(9)	12	(1)	15	(1)	47	(3)
	20	(1)	16	(15)	27	(5)		
	35	(13)	47	(5)	0	(3)		
	22	(15)	0	(3)	27	(10)		
	13	(15)						
	39	(5)						
	18	(6)						
Median	20		14		27		35	
Mean	22		20		24		44	

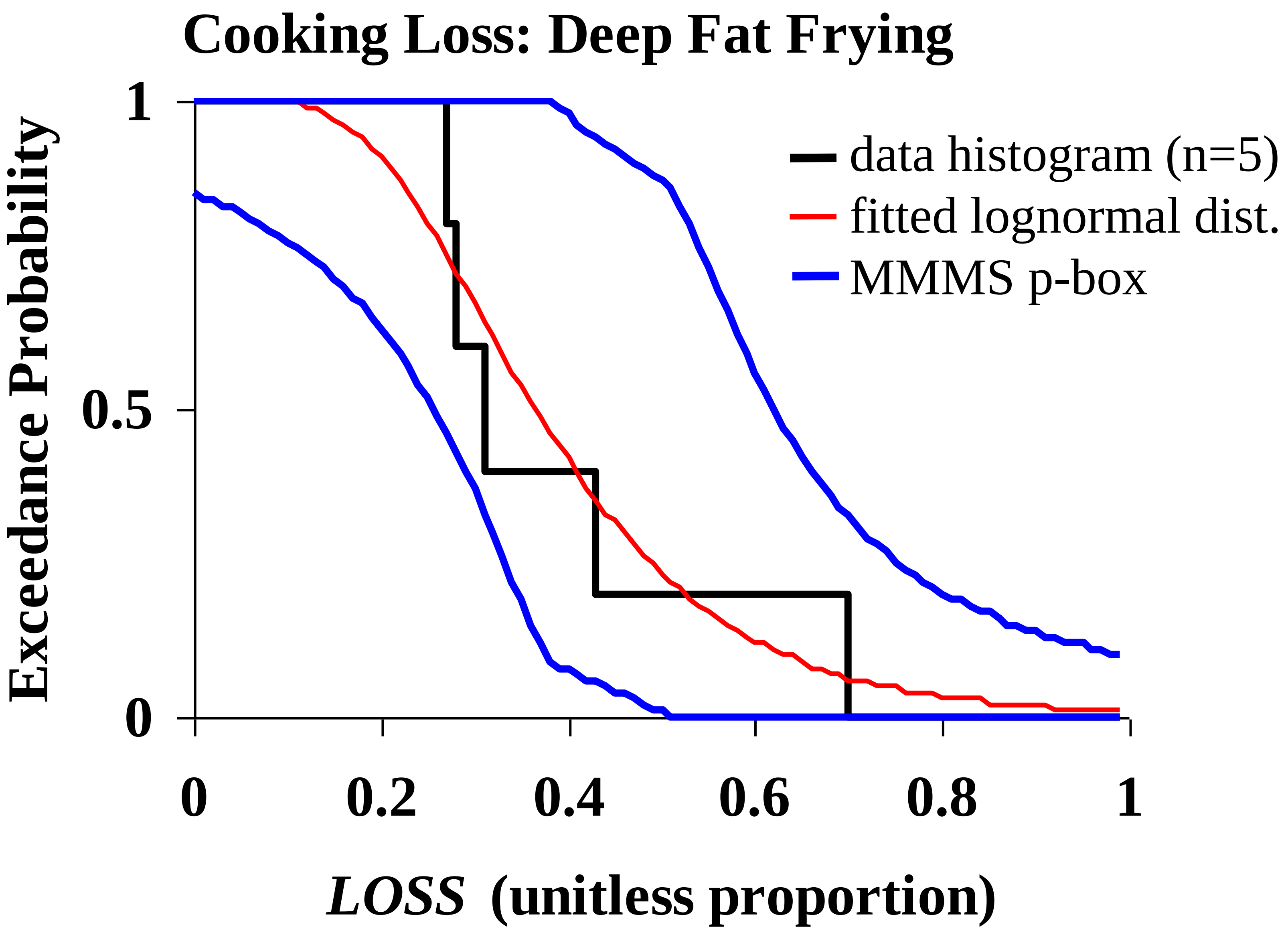


Cooking Loss: Baking
Exceedance Probability
1
0.5
0
0
0.2
0.4
0.6
0.8
1
LOSS (unitless proportion)
data histogram (n=11)
fitted lognormal distribution
MMMS p-box



Cooking Loss:Broiling
Exceedance Probability
1
0.5
0
0
0.2
0.4
0.6
0.8
1
LOSS (unitless proportion)
data histogram (n=8)
fitted lognormal distribution
MMMS p-box


Cooking Loss: Pan Frying
Exceedance Probability
1
0.5
0
0
0.2
0.4
0.6
0.8
1
data histogram (n=8)
fitted lognormal distribution
MMMS p-box
LOSS (unitless proportion)


Cooking Loss: Deep Fat Frying
Exceedance Probability
1
0.5
0
data histogram (n=5)
fitted lognormal dist.
MMMS p-box
0
0.2
0.4
0.6
0.8
1
LOSS (unitless proportion)

Cooking method weights

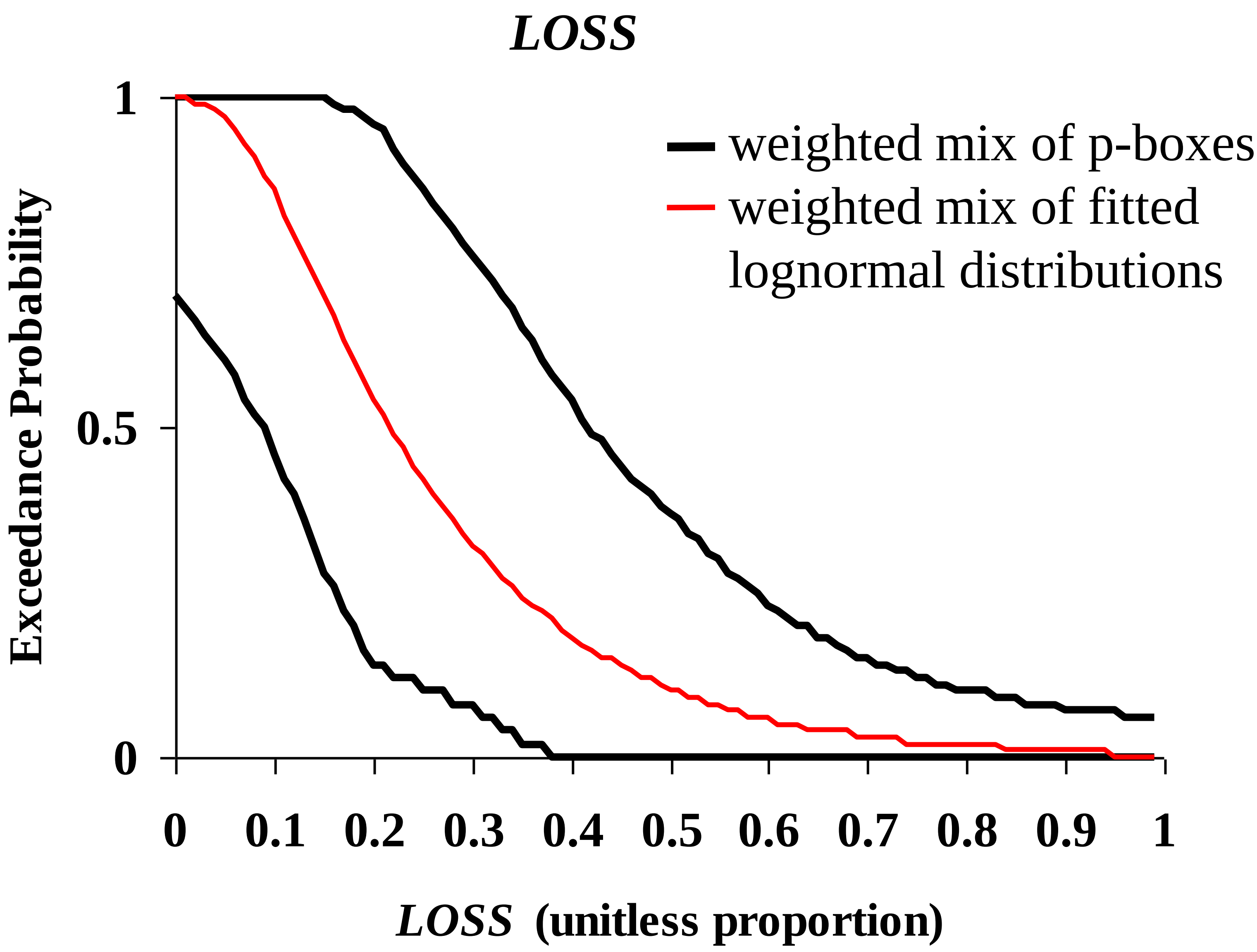
Data

Method	%
Frying (pan & deep fat)	62.1
Baking	17.9
Broiling	16.4
Soup	2.0
Poaching	0.9
Microwaving	0.9
Raw	0.6
Boiling	0.2

96.4%

→ Decision

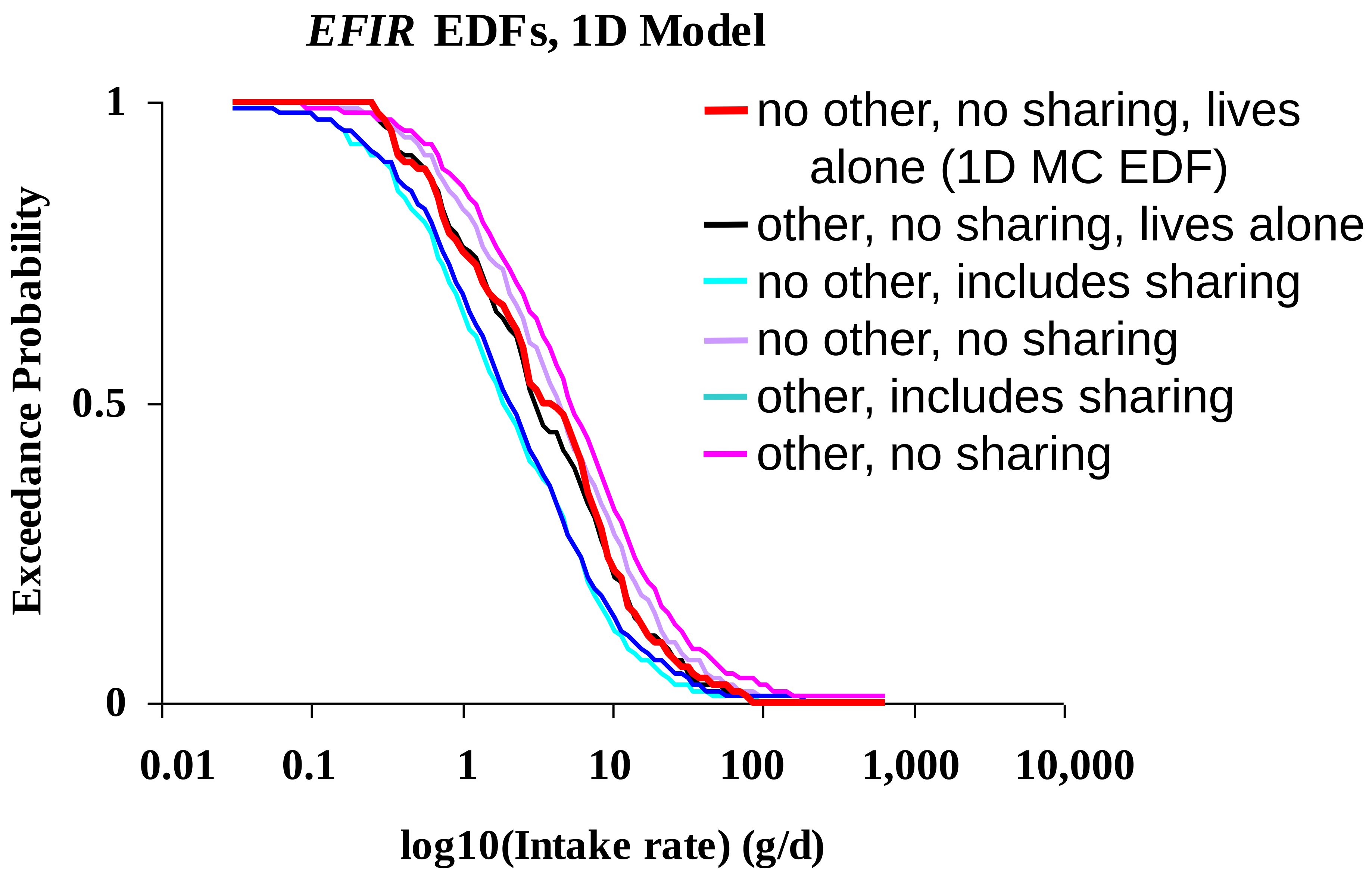
Method	%
Pan frying	40
Deep fat frying	20
Baking	20
Broiling	20



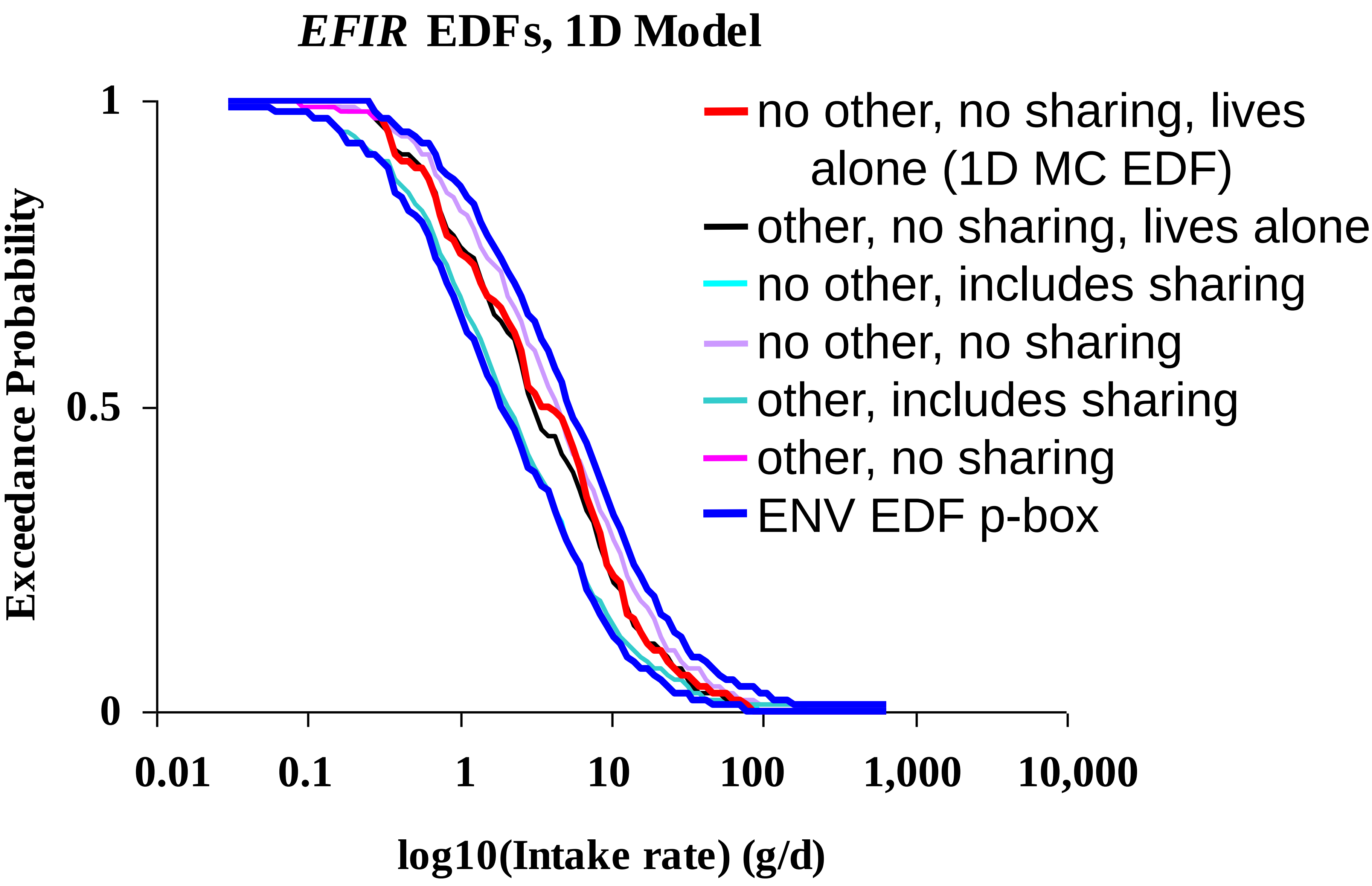
LOSS
weighted mix of p-boxes
weighted mix of fitted lognormal distributions
Exceedance Probability
1
0.5
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
LOSS (unitless proportion)

1D model intake rate

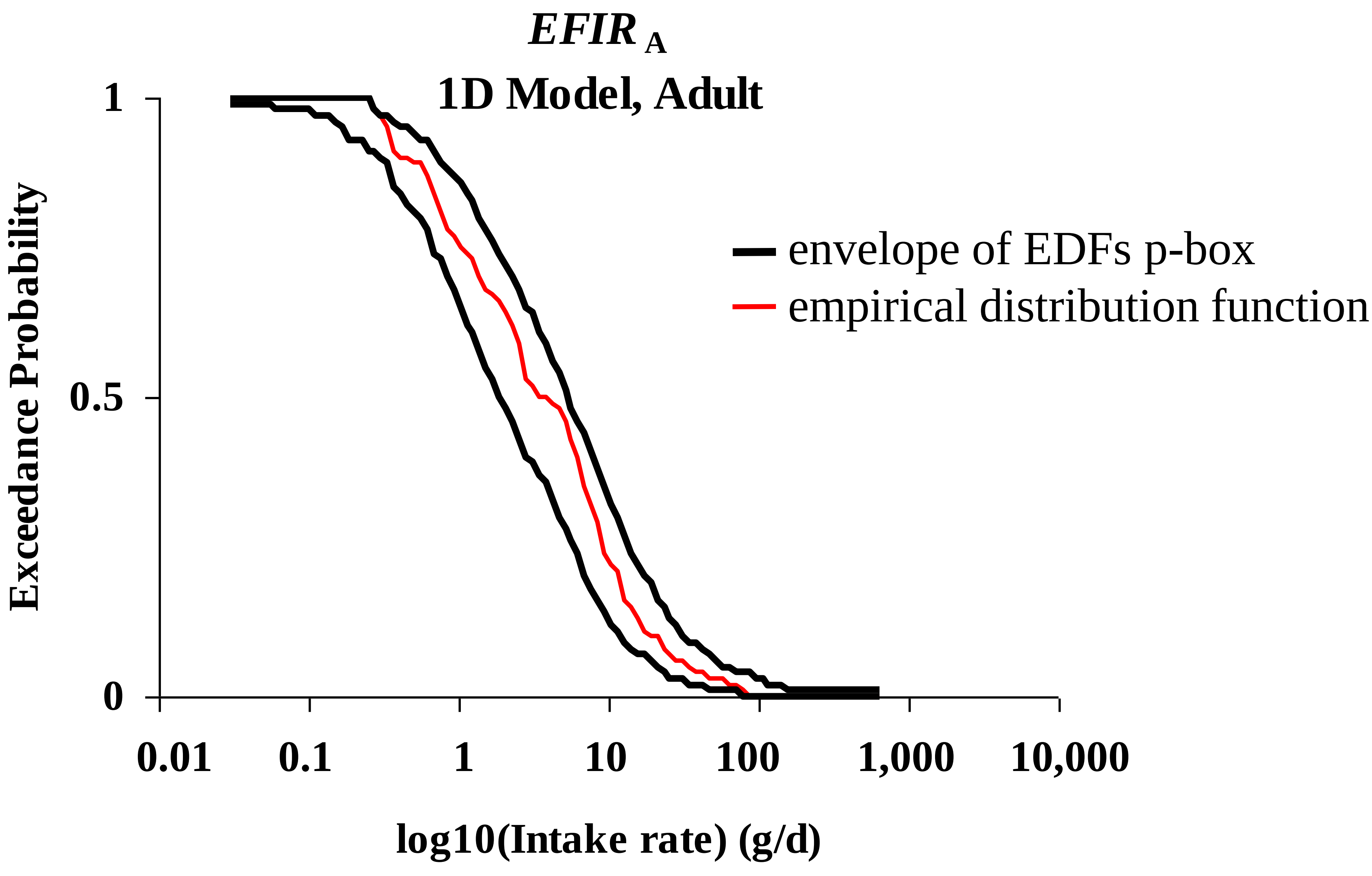
- 1D model uses data from the Maine Angler Survey to model intake rate.
- The survey measured grams of game fish consumed per household per day (g/d).
- The data was stratified:
  - Did you consume fish from “other” sources?
  - Did you eat shared fish?
  - Do you live alone?
    - If not, intake rate divided by number in household.



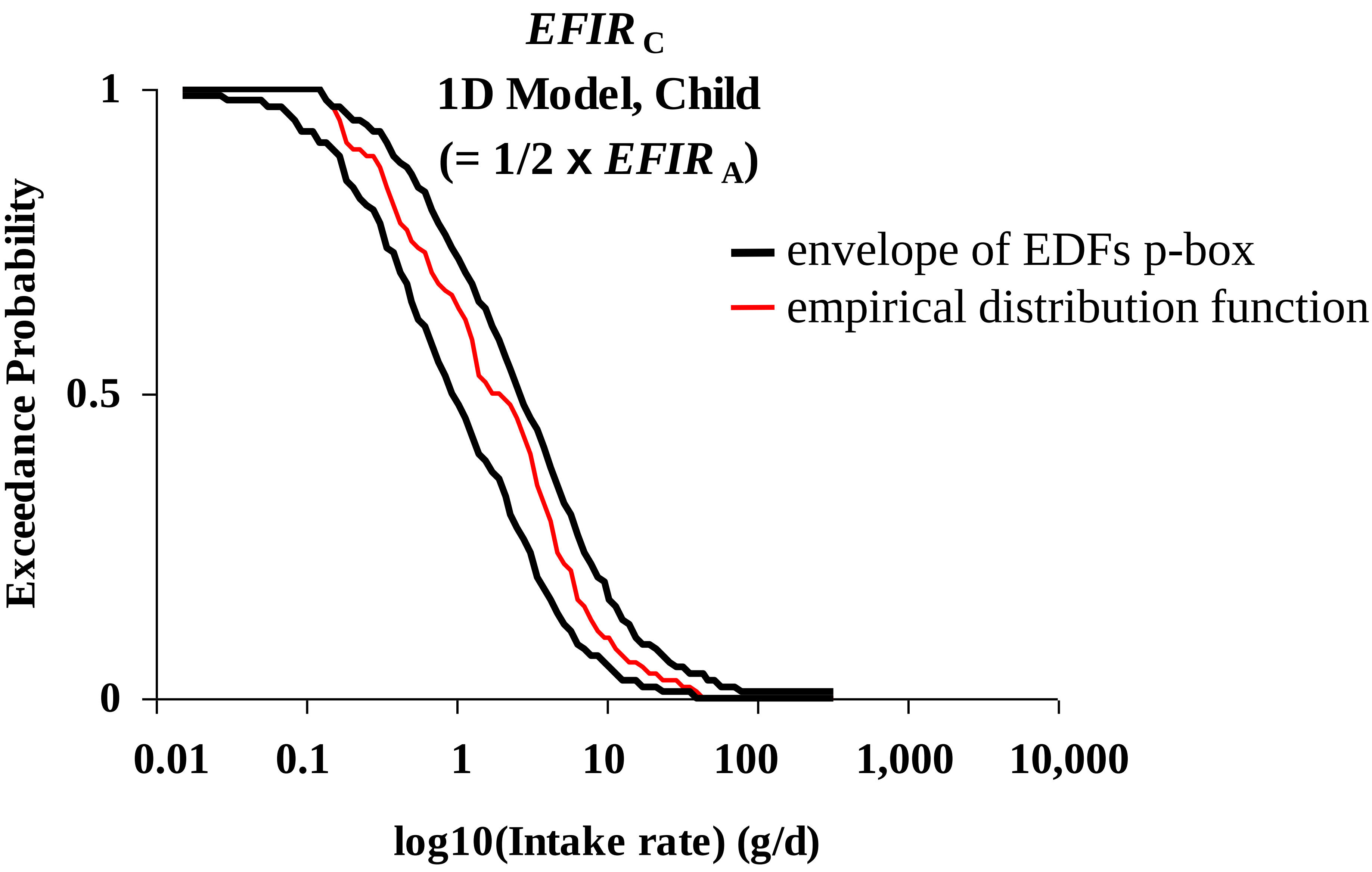
EFIR EDFs, 1D Model
Exceedance Probability
1
0.5
0
0.01
0.1
1
10
100
1,000
10,000
log10(Intake rate) (g/d)
no other, no sharing, lives alone (1D MC EDF)
other, no sharing, lives alone
no other, includes sharing
no other, no sharing
other, includes sharing
other, no sharing



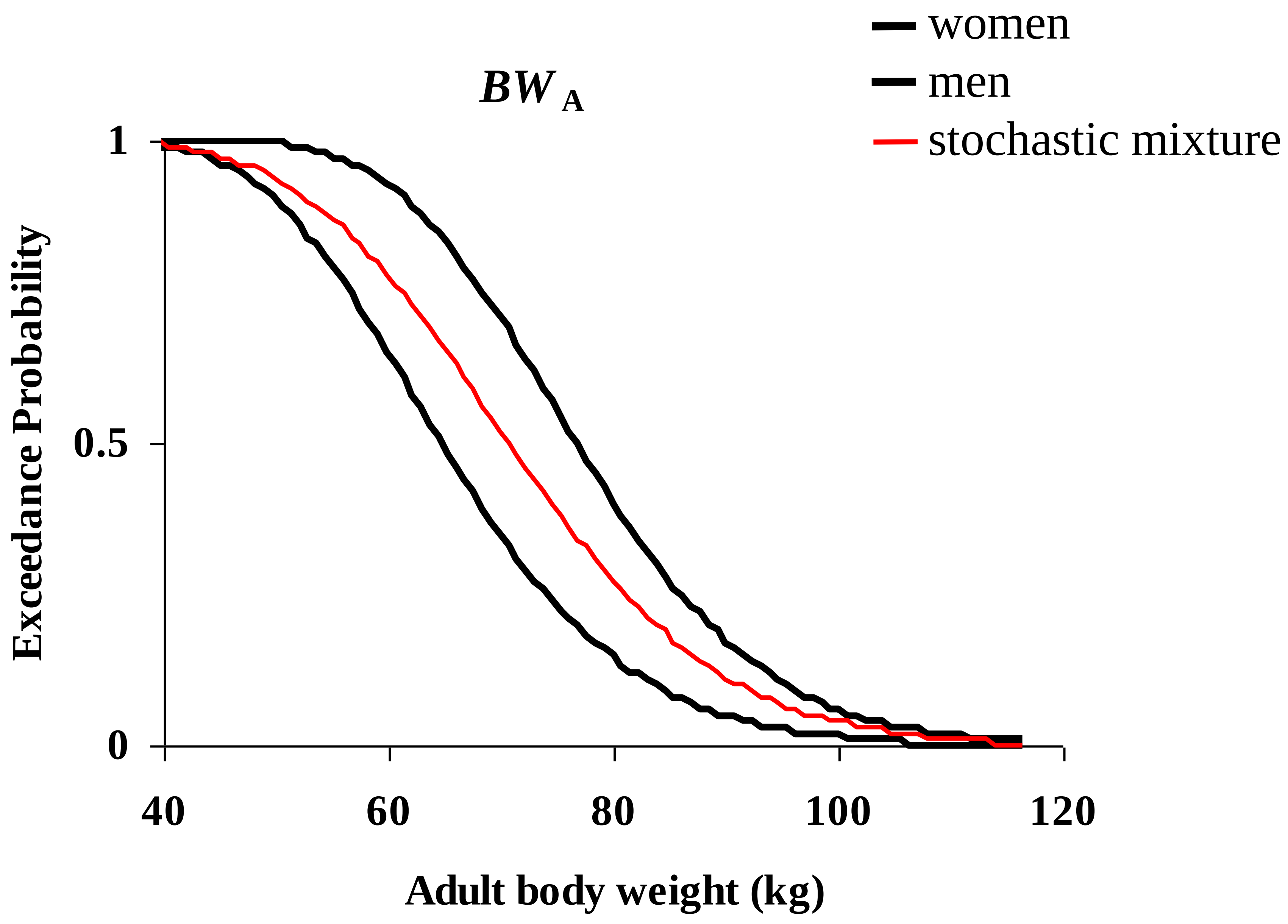
EFIR EDFs, 1D Model
Exceedance Probability
1
0.5
0
0.01
0.1
1
10
100
1,000
10,000
log10(Intake rate) (g/d)
no other, no sharing, lives alone (1D MC EDF)
other, no sharing, lives alone
no other, includes sharing
no other, no sharing
other, includes sharing
other, no sharing
ENV EDF p-box



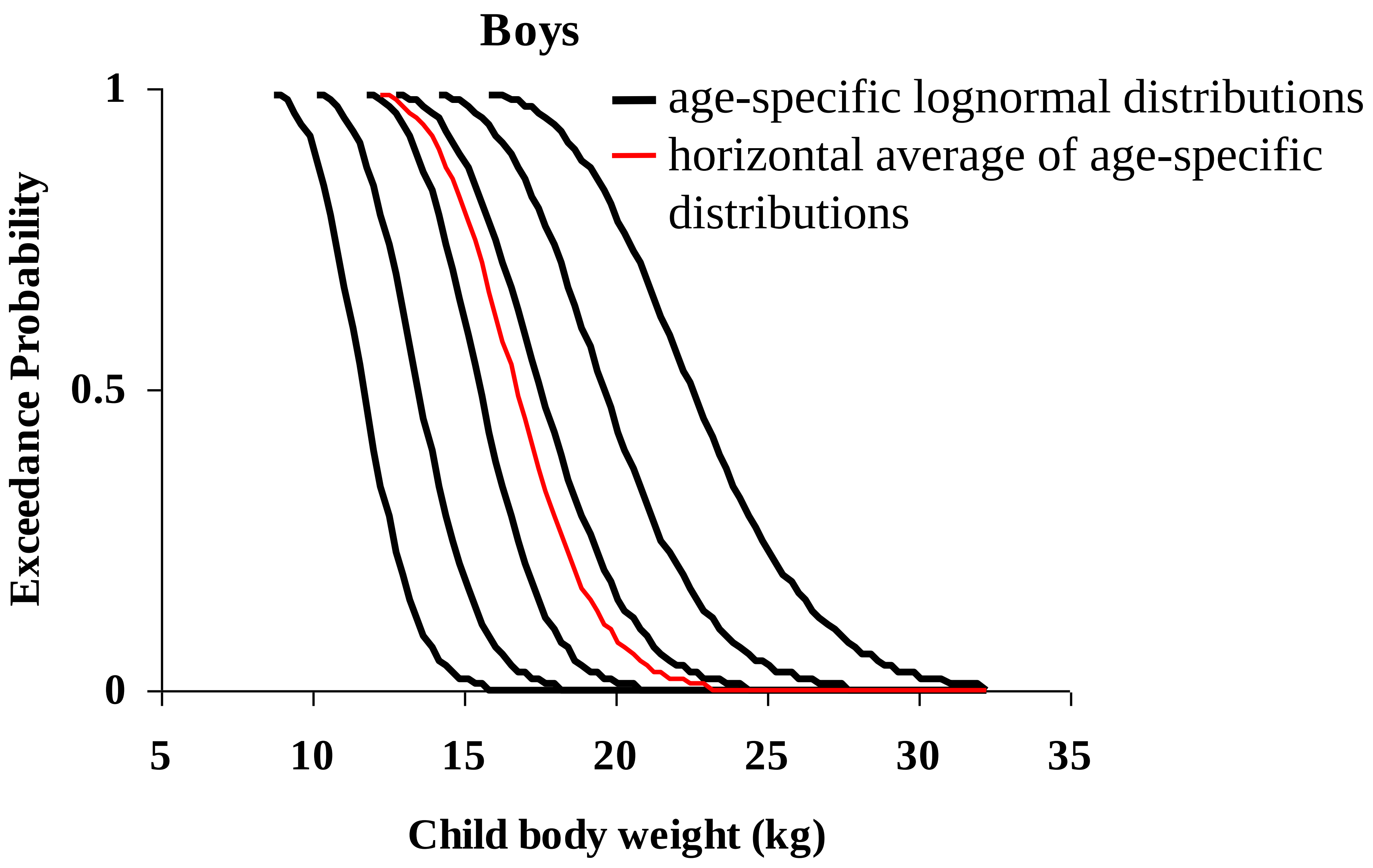
EFIR A
1D Model, Adult
Exceedance Probability
1
0.5
0
envelope of EDFs p-box
empirical distribution function
0.01
0.1
1
10
100
1,000
10,000
log10(Intake rate) (g/d)



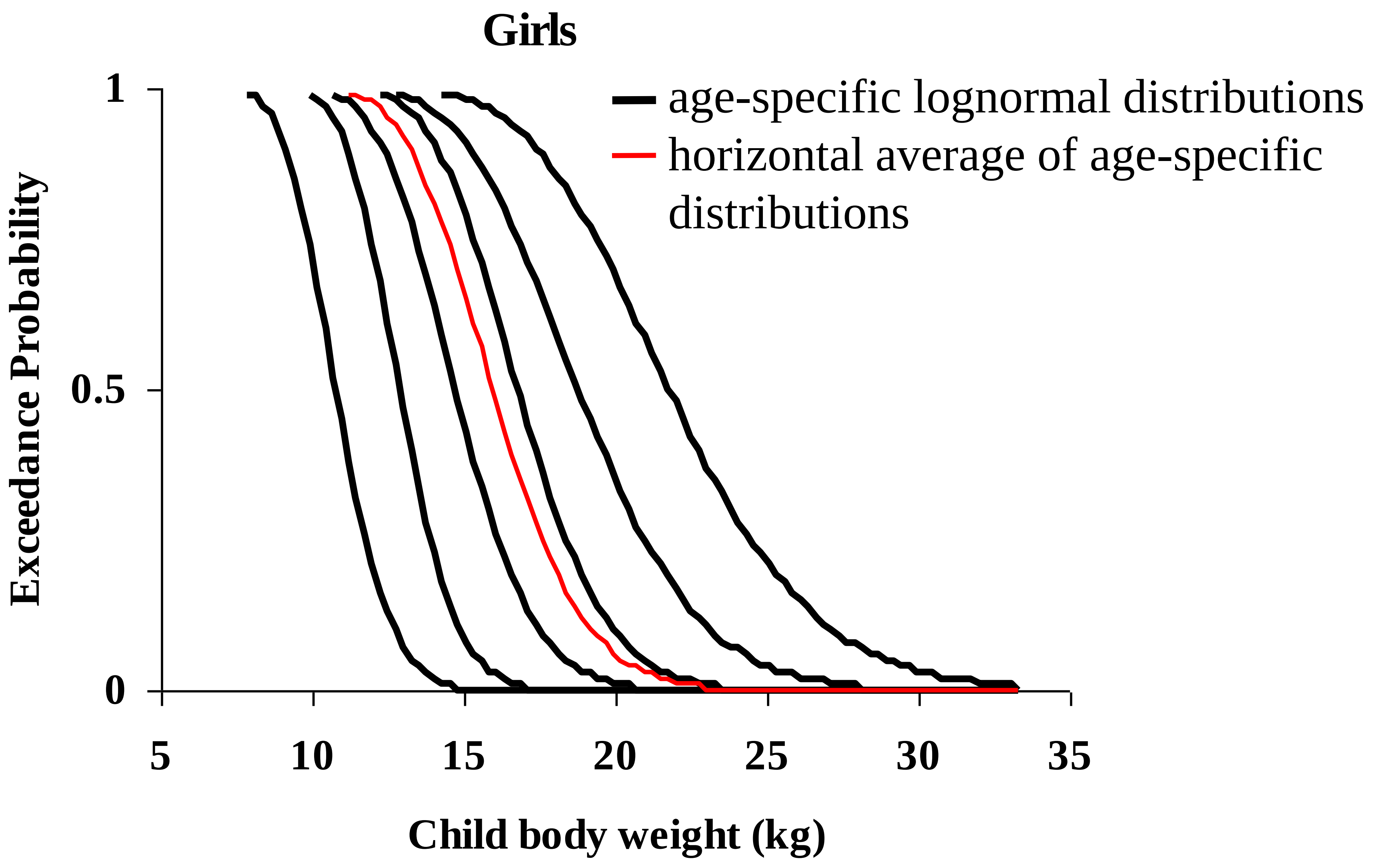
EFIR C
1D Model, Child
(= 1/2 x EFIR A)
envelope of EDFs p-box
empirical distribution function
Exceedance Probability
1
0.5
0
0.01
0.1
1
10
100
1,000
10,000
log10(Intake rate) (g/d)

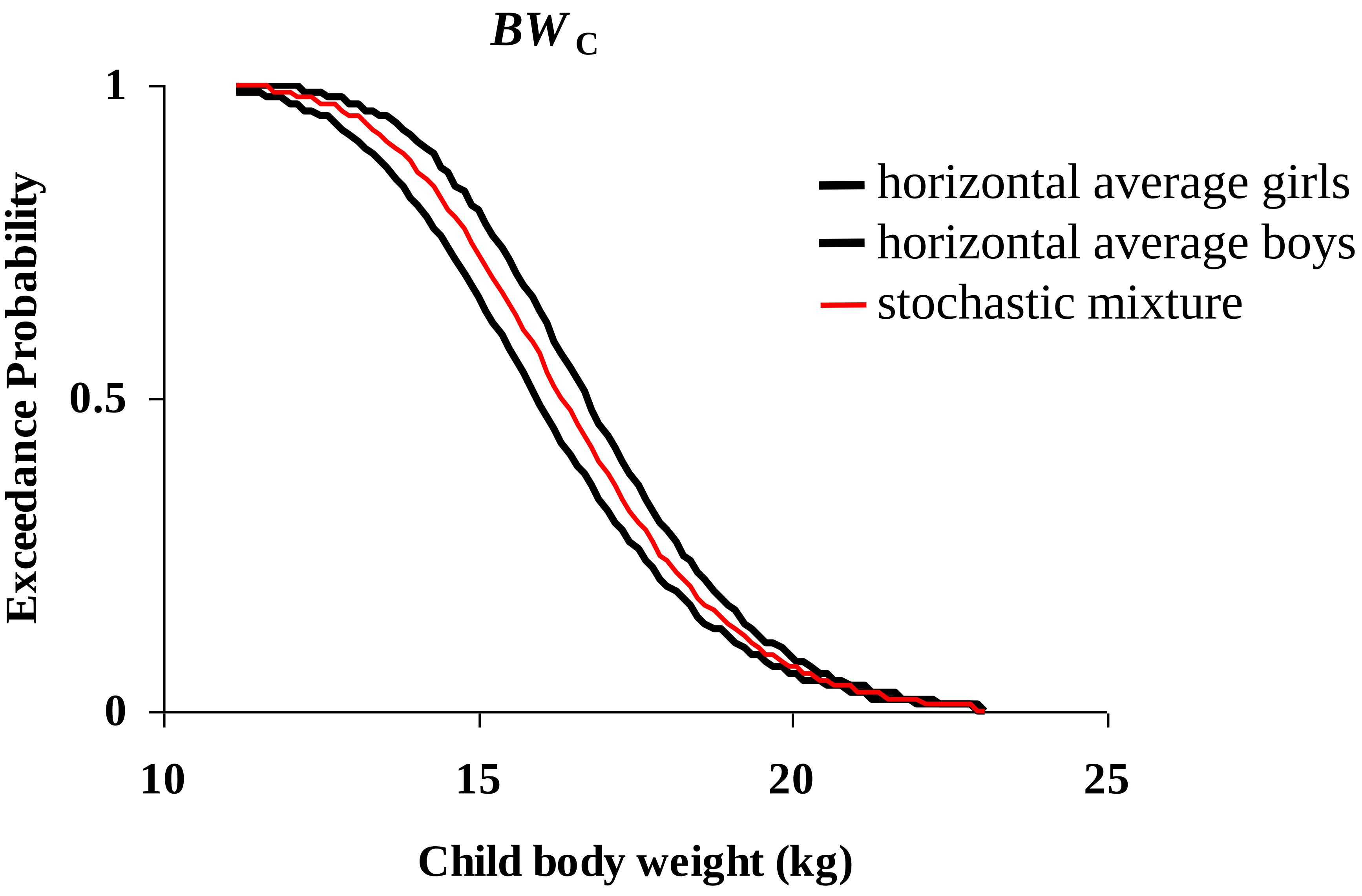


BW_A
women
men
stochastic mixture
Exceedance Probability
1
0.5
0
40
60
80
100
120
Adult body weight (kg)

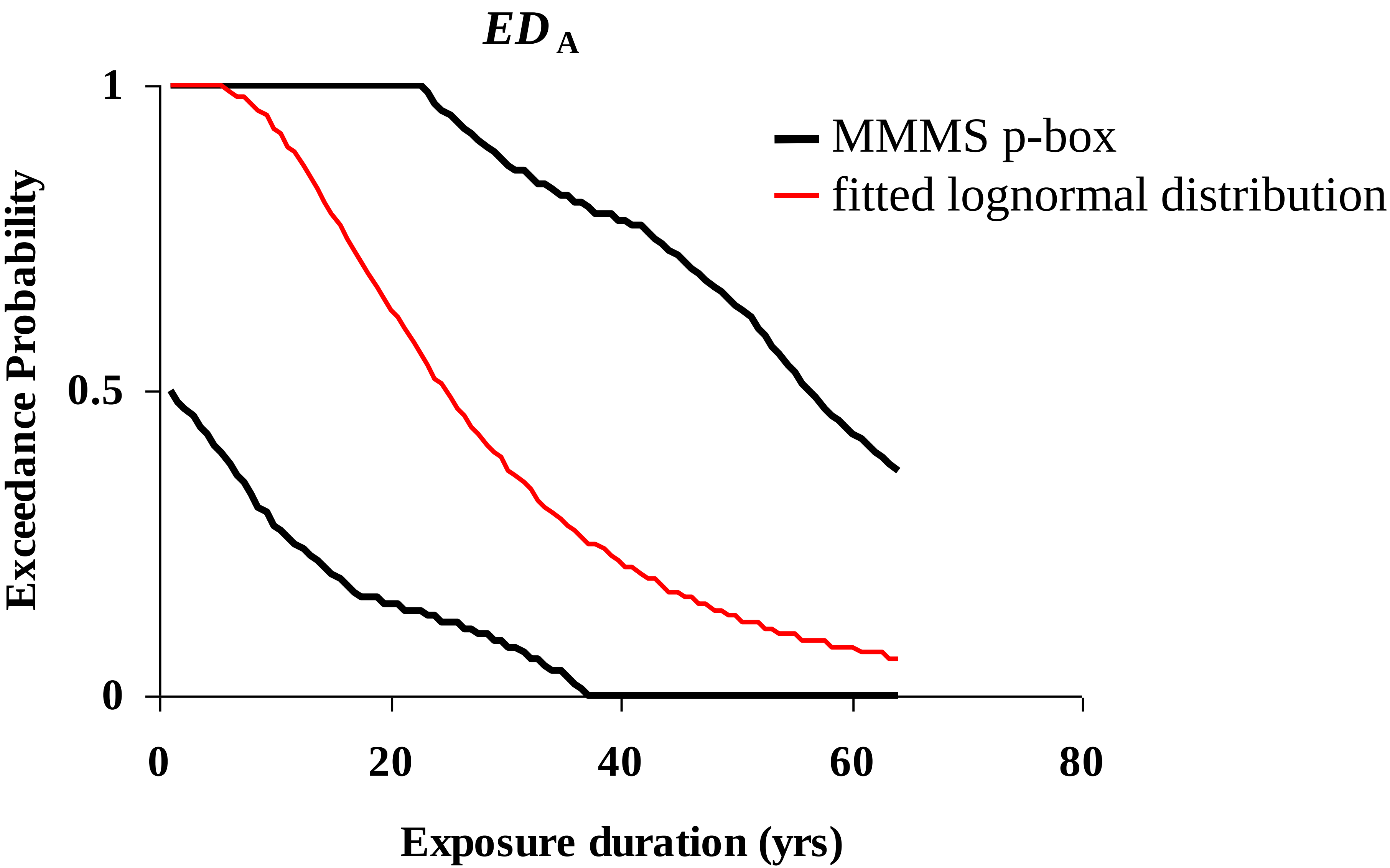


Boys
age-specific lognormal distributions
horizontal average of age-specific distributions
Exceedance Probability
1
0.5
0
5
10
15
20
25
30
35
Child body weight (kg)

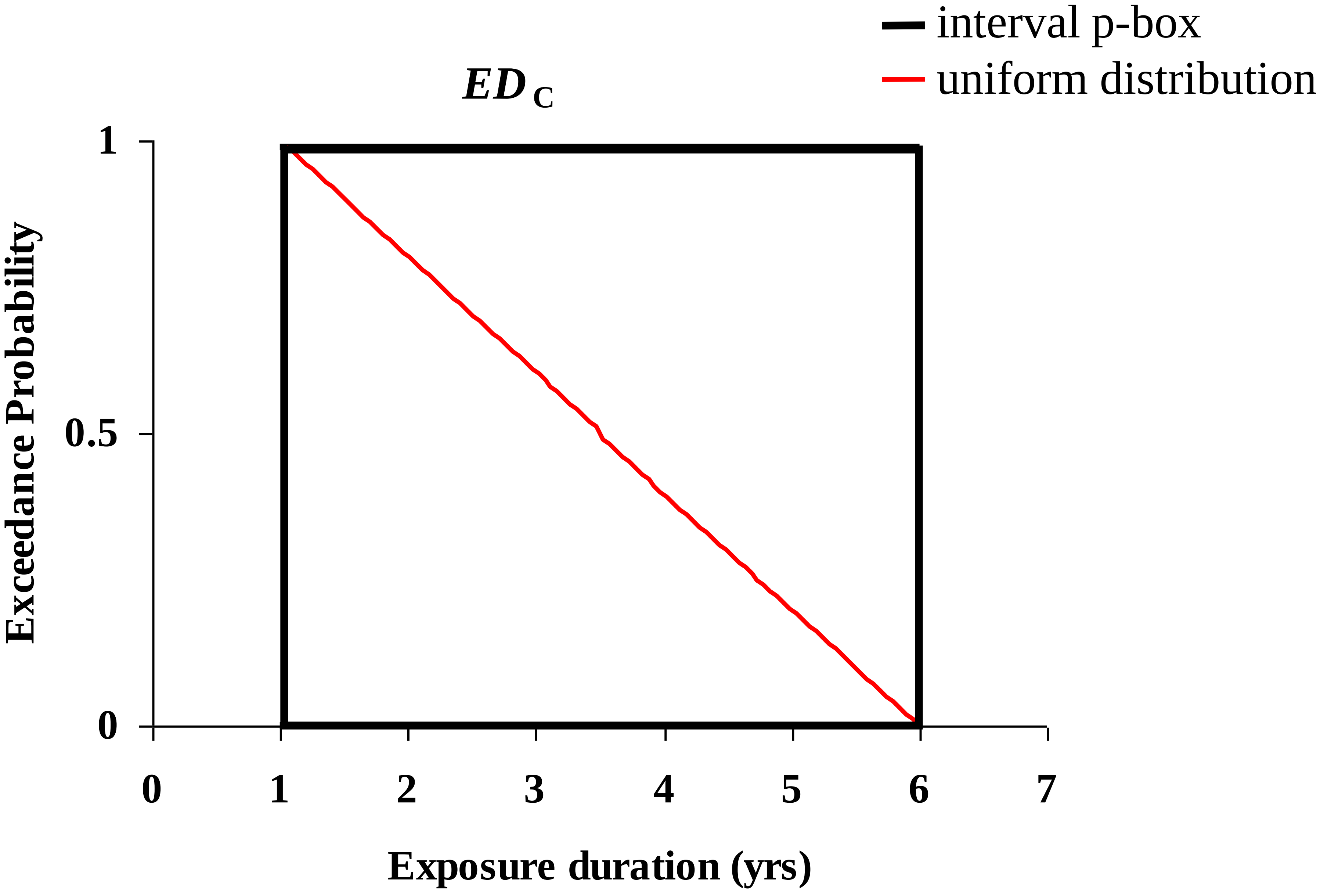

Girls
age-specific lognormal distributions
horizontal average of age-specific distributions
Exceedance Probability
1
0.5
0
5
10
15
20
25
30
35
Child body weight (kg)



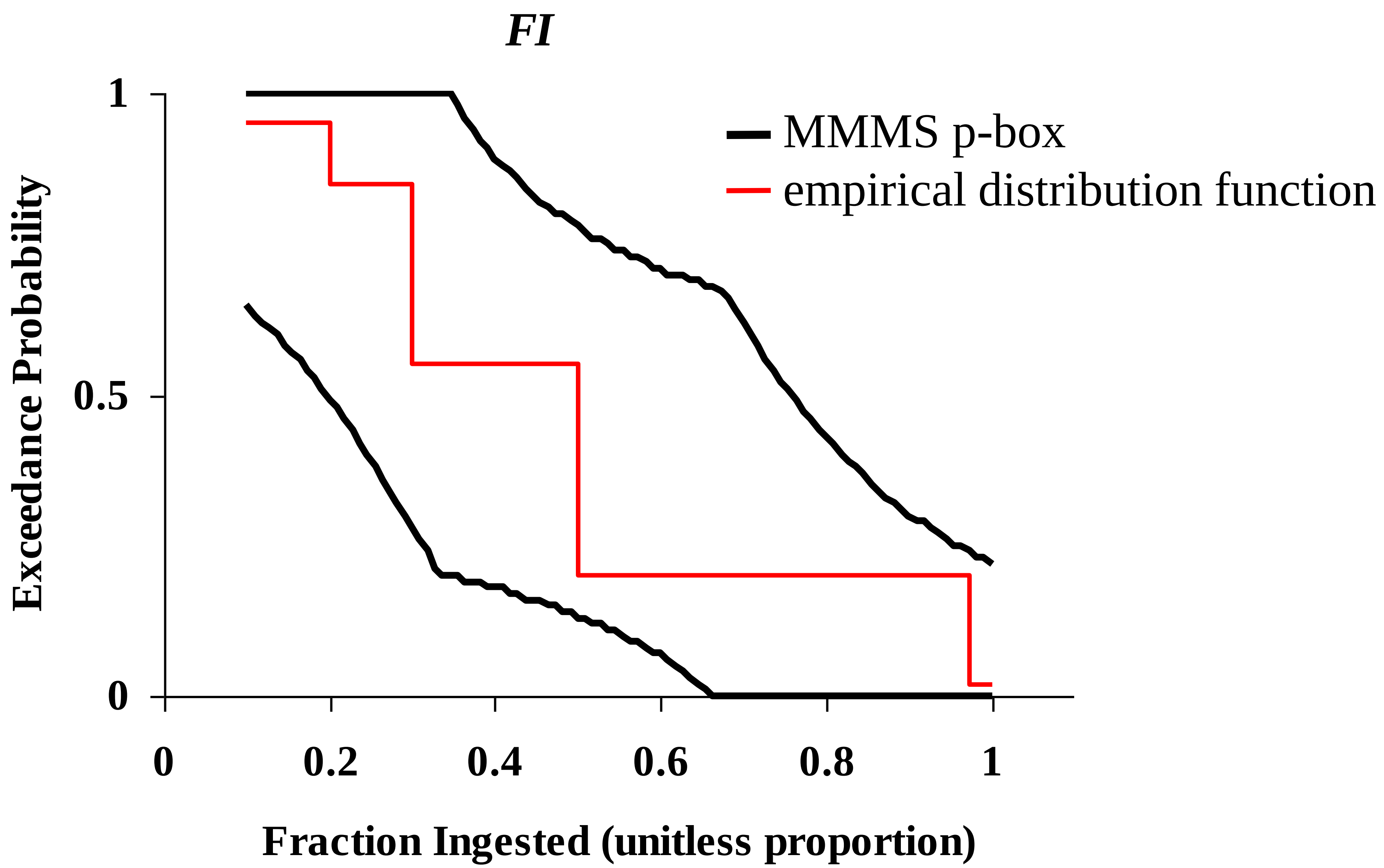
BW C
1
0.5
0
10
15
20
25
Exceedance Probability
Child body weight (kg)
horizontal average girls
horizontal average boys
stochastic mixture



ED_A
MMMS p-box
fitted lognormal distribution
Exceedance Probability
1
0.5
0
0
20
40
60
80
Exposure duration (yrs)



interval p-box
uniform distribution
ED_C
1
0.5
0
Exceedance Probability
0
1
2
3
4
5
6
7
Exposure duration (yrs)



FI
MMMS p-box
empirical distribution function
Exceedance Probability
1
0.5
0
0
0.2
0.4
0.6
0.8
1
Fraction Ingested (unitless proportion)

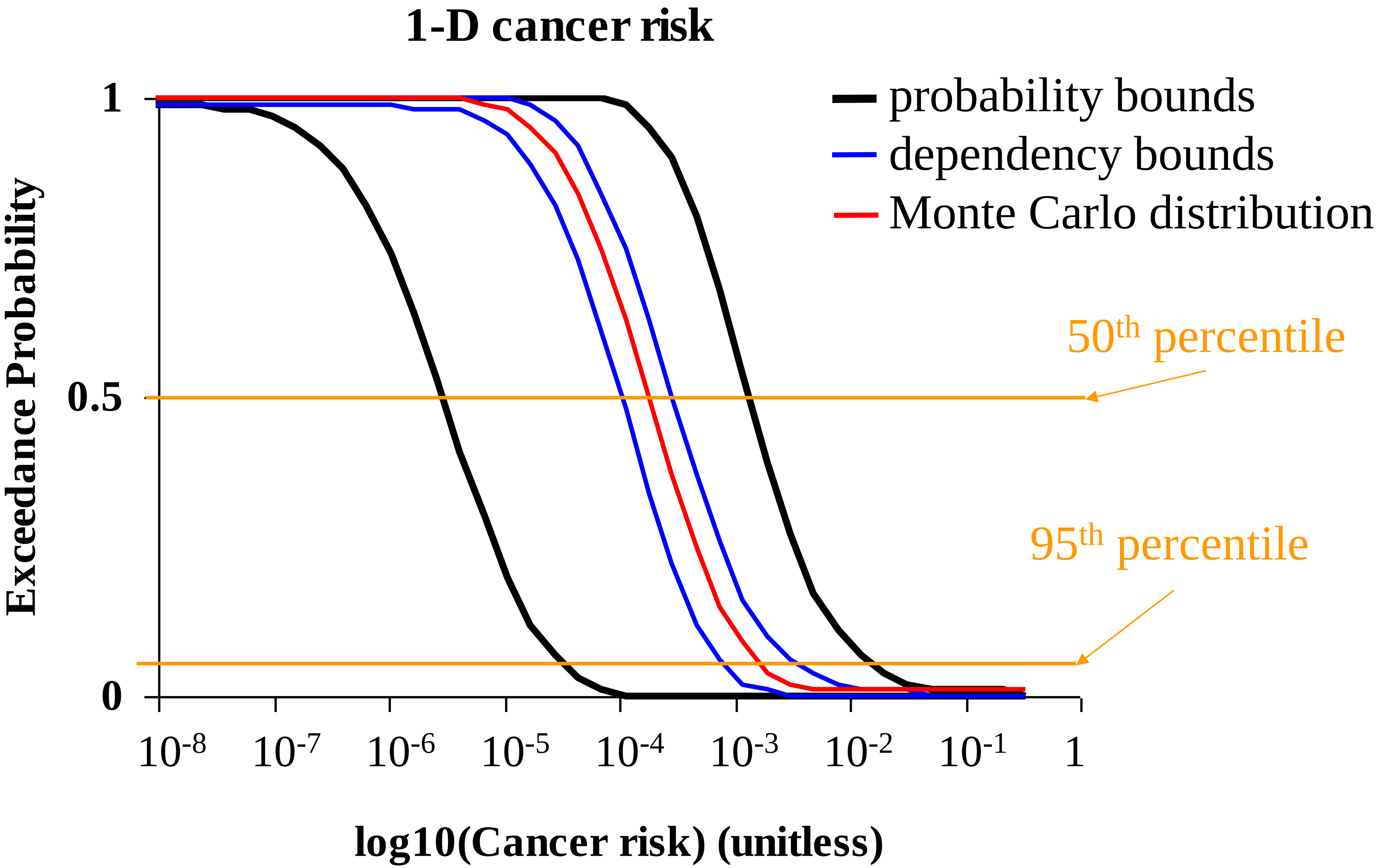
Dependency bounds analysis

	C	LOSS	IR	EF	FI	ED
LOSS	I					
IR	I	I				
EF	I	I	?			
FI	I	I	I	I		
ED	I	I	?	I	I	
BW	I	I	?	I	I	?

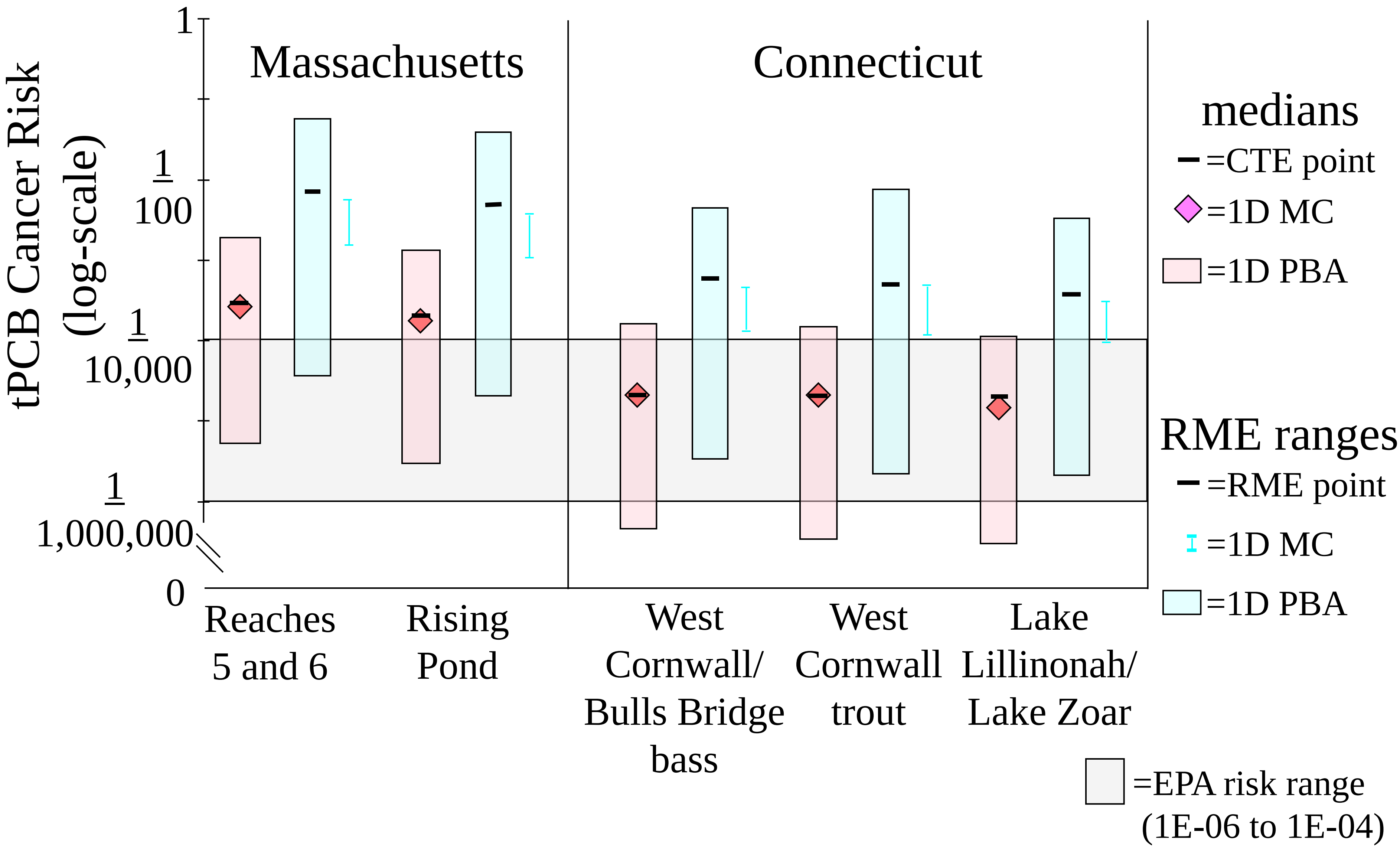
"I" indicates independence assumed

"?" indicates possible dependency relationship

Risk at Rising Pond

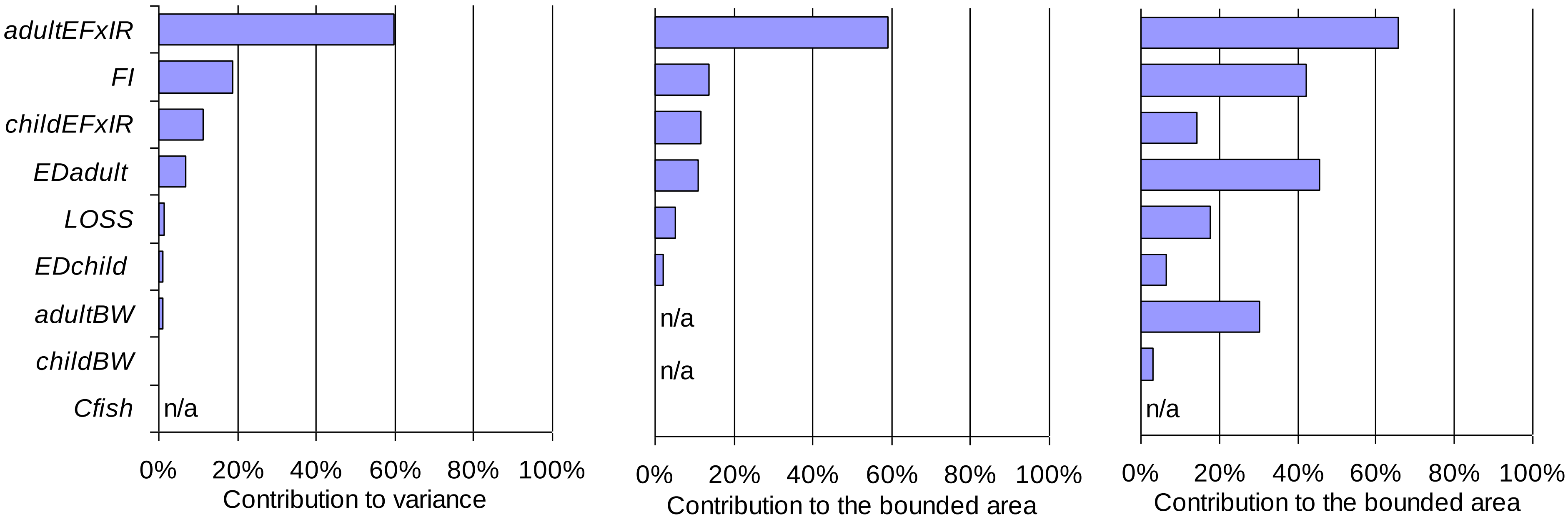


Cancer risk at five sites



Sensitivity: cancer model

1D MC Coefficient of determination ratio	1D PBA Remove uncertainty	1D PBA Remove uncertainty and variability



Case study references

Weston Solutions, Inc. 2005. Probabilistic Risk Assessment. Section 6 in *Human Health Risk Assessment, GE/Housatonic River Site, Rest of River, Volume 4, Appendix C, Fish and Waterfowl Consumption Risk Assessment*, prepared for US Army Corps of Engineers and US EPA. February 11, pp. 6-1 – 6-129. Available online at http://www.epa.gov/region1/ge/thesite/restofriver/reports/hhra_219190/219190_HHRA_Vol4_FW.pdf.

Weston Solutions, Inc. 2005. Probability Bounds Analysis Approach. Attachment 5 in *Human Health Risk Assessment, GE/Housatonic River Site, Rest of River, Volume 1*, prepared by Weston Solutions, Inc. for US Army Corps of Engineers and US EPA. February 11. Available online at http://www.epa.gov/region1/ge/thesite/restofriver/reports/hhra_219190/219190_HHRA_FNL_VolumeTOC.pdf.

Cooking Loss Data Sources:

(1) Armbruster 1987; *Journal of Food Safety* 8:235-243.

(2) Armbruster 1989; *Journal of Food Safety* 9:235-244.

(3) Moya et al. 1998; *Bull. Environ. Contam. Toxicol* 60:845-851.

(4) Puffer and Gossett 1983; *Bull. Environ. Contam. Toxicol.* 30:65-73.

(5) Salama et al. 1998; *J. Agric. Food Chem.* 46:1359-1362.

(6) Schechter et al. 1998; *Chemosphere* 37:1723-1730.

(7) Skea eat al. 1979; *J. Great Lakes Res., Internat. Assoc. Great Lakes Res.* 5 (2):153-159.

(8) Smith et al. 1973; *J. Fish. Res Board Can.* 30:702-706.

(9) Trotter et al. 1989; *J. Assoc. Off. Anal. Chem.* 75:501-503.

(10) Wang et al. 2000; *Human Exposure Posters – Organohalogen Compounds. Volume 48.* Division of Environmental Health and Risk Management, University of Birmingham, Birmingham, UK.

(11) Zabik et al. 1979; *Bull. Environ. Contam. Toxicol.* 21:136-143.

(12) Zabik et al. 1982; *Bull. Environ. Contam. Toxicol.* 28:710-715.

(13) Zabik et al. 1995a; *Bull. Environ. Contam. Toxicol.* 54:396-402.

(14) Zabik et al. 1995b; *J. Agric. Food Chem.* 43:993-1001.

(15) Zabik et al. 1996; Pesticide residues, PCBs and PAHs in baked, charbroiled, salt boiled and smoked Great Lakes lake trout.

Cooking Loss weights and EFIR data sources:

ChemRisk. 1992. *Consumption of Freshwater Fish by Maine Anglers*. 24 July 1992.

ChemRisk. 1995. *Evaluating the Impact of Cooking Processes on the Level of PCBs in Fish*. Prepared for General Electric Company. January 1995.

Body weight distribution data sources:

Brainard J and Burmaster D.E. 1992. Bivariate distributions for height and weight of men and women in the United States. *Risk Analysis* 12: 267-275.

Burmaster, D.E. and E.A.C. Crouch 1997. Lognormal distributions for body weight as a function of age for males and females in the United States, 1976-1980. *Risk Analysis* 17(4):499-505.

ED data source:

MDPH (Massachusetts Department of Public Health). 2001. Letter from Suzanne K. Condon, Assistant Commissioner of the Bureau of Environmental Health Assessment to Bryan Olson, U.S. Environmental Protection Agency, Region I. Tables with Hunting Information for Individual Family Members Who Reported Hunting Birds from the HRA, PCB Exposure Assessment Study, Volunteer Study, and Hotline Study and Calls from Individuals Concerned about Hunting after Hearing about the PCB Duck Advisory. 21 August 2001.

FI data sources:

Barry, T.J. 1988. *An Angler Survey and Economic Study of the Housatonic River Fishery Resource.* Connecticut Department of Environmental Protection, Bureau of Fisheries, Hartford. Final Report.

ChemRisk. 1992. *Consumption of Freshwater Fish by Maine Anglers*. 24 July 1992.

Ebert, E.S., S.H. Su, T.J. Barry, M.N. Gray, and N.W. Harrington. 1996. Estimated rates of fish consumption by anglers participating in the Connecticut Housatonic River Creel Survey. *North American Journal of Fisheries Management* 16:81-89.

Conclusions

Take-home messages

- Monte Carlo will always be useful, like Euclidean geometry
- Using bounding, you don't have to pretend you know a lot to get *quantitative* answers
- Approximation and bounding are often complementary
- Paying attention to measurement imprecision, censoring, and missingness requires new approaches
- Humans treat incertitude and variability differently and risk analysts should respect this

Untenable assumptions

- Distributions normal
- Uncertainties are small
- Sources of variation are independent
- Uncertainties can cancel each other out
- Linearized models good enough
- Most of the physics is known and modeled

Need ways to relax assumptions

- Possibly large, non-normal uncertainties
- Non-independent, or *unknown* dependencies
- Uncertainties that may not cancel
- Arbitrary mathematical operations
- Model uncertainty

Everyone makes assumptions

- But not all sets of assumptions are equal!

Point value
Interval range
Entire real line

Linear function
Monotonic function
Any function

Normal distribution
Unimodal distribution
Any distribution

Independence
Known correlation
Any dependence

- Like to discharge unwarranted assumptions

"Certainties lead to doubt; doubts lead to certainty"

What is known empirically	**Monte Carlo analysis**	**Probability bounds analysis**
Know only range of variable	Assume uniform distribution	Assume interval
Know some constraints about random variable	Select largest entropy distribution from all thus constrained distributions	Form envelope around class of distributions matching constraints
Uncertainty about distribution family or shape	Repeat analysis for other plausible distribution shapes	Form distribution-free p-box as envelope of all plausible distributions
Sample data	Form empirical distribution function (EDF)	Form nonparametric c-box or KS confidence limits around EDF
Variable follows known marginal distribution	Sample from particular distribution	Use particular distribution
Measurement uncertainty	Ignore it (usually), or perform sensitivity analysis	Express it in intervals and incorporate it into analysis

What is known empirically	Monte Carlo analysis	Probability bounds analysis
Non-detects	Replace non-detect with ½DL (detection limit)	Replace non-detect with interval [0, DL]
Know variables are independent	Assume independence	Assume (random-set) independence
Know magnitude of correlation	Simulate correlation from particular (but usually arbitrary) copula	Bound result from possible copulas with correlation, or use known copula
Know only the general sign (+ or –) of dependence	Assume some correlation of appropriate sign, or repeat analysis for different correlations	Bound result assuming only the sign of the dependence and specific or all possible copulas
Do not know the nature of the dependence	Assume independence (usually), or repeat analysis for different correlations	Bound result for all possible dependencies (Fréchet case)
Model uncertainty	Form stochastic mixture (vertical average) of distribution functions	Form envelope of distribution functions

Cheat sheet for R libraries

- plot, lines, show, summary
- mean, sd, var, median, quantile, left, right, prob, cut, percentile, iqr, random, range
- exp, log, sqrt, abs, round, trunc, ceiling, floor, sign, sin, cos, tan, asin, acos, atan, atan2, reciprocate, negate, +, -, *, /, pmin, pmax, ^, and, or, not, mixture, smin, smax

sra.r

normal, etc.

histogram

quantiles

MM<tab><tab>

ME<tab><tab>

ML<tab><tab>

pba.r

normal, etc.

histogram

quantiles

pointlist

MM<tab><tab>

ME<tab><tab>

ML<tab><tab>

CB<tab><tab>

NV<tab><tab>

Supported named distributions

bernoulli, beta, binomial, cauchy, chi, chisquared, delta, dirac, discreteuniform, exponential, exponentialpower, extremevalue, f, fishersnedecor, fishertippett, fisk, frechet, gamma, gaussian, geometric, generalizedextremevalue, generalizedpareto, gumbel, histogram, inversegamma, laplace, logistic, loglogistic, lognormal, logtriangular, loguniform, negativebinomial, normal, pareto, pascal, powerfunction, poisson, quantiles, rayleigh, reciprocal, shiftedloglogistic, skewnormal, student, trapezoidal, triangular, uniform, Weibull

Nonparametric p-boxes

maxmean, minmax, minmaxmean, minmean, meanstd, meanvar, minmaxmode, minmaxmedian, minmaxmedianismode, minmaxpercentile, minmaxmeanismedian, minmaxmeanismode, mmms, mmmv, posmeanstd, symmeanstd, uniminmax, unimmmv, unimmms

Functions histogram, quantiles differ

```
sra.r
d = c(0,1,2,4,7,12)
k = length(d) - 1
p = 0:k/k
q = quantiles(d,p)
h =histogram(d)
p; blue(h)

d = runif(10)
ml = MLlaplace(x)
mm = MMlaplace(x)
edf(d)
red(ml)
blue(mm)
```

```
pba.r
d = c(0,1,2,4,7,12)
k = length(d) - 1
p = 0:k/k
a = quantiles(d,p)
b = pointlist(d,p)
a; red(b)

d = c(3,4,2,2,3,3,4,5,5,3,2,3)
e = histogram(d)
f = histogram(d,conf=0)
e
black(e)
green(f)
```

Synopsis: Monte Carlo simulation

- How?
  - replace point estimates with distributions
  - repeatedly sample from each distribution and compute
  - tally answers in histogram
- Why?
  - simple to implement
  - fairly simple to explain
  - summarizes entire distribution of risk
- Why not?
  - requires a lot of empirical information (or guesses)
  - routine assumptions may be “non-protective”

Synopsis: probability bounds analysis

How?

- specify what you are sure about
- establish bounds on probability distributions
- pick dependencies (no assumption, indep., perfect, etc.)

Why?

- account for uncertainty better than maximum entropy, etc.
- puts bounds on Monte Carlo results
- bounds get narrower with better empirical information

Why not?

- does not yield second-order probabilities
- best-possible results can sometimes be expensive to compute

Synopsis: imprecise probabilities

How?

- avoid sure loss, $\underline{P}(A) \leq \overline{P}(A)$
- be coherent, $\underline{P}(A) + \underline{P}(B) \leq \underline{P}(A \cup B)$
- use natural extension (mathematical programming) to find consequences

Why?

- most expressive language for uncertainty of *all kinds*
- can provide expectations and conditional probabilities
- provides best possible results that do not lose information

Why not?

- requires mathematical programming
- can strain mathematical ability of the analyst

Software for Monte Carlo

- R library `sra.r`, and the R language itself
- Crystal Ball
- @Risk by Palisade
- Analytic by Lumina
- GoldSim, Solver, ModelRisk, etc.
- Monte Carlo methods are embedded in many special-purpose simulation tools (e.g., RAMAS library for ecological risk analysis)

Software for probability bounds

- R libraries pbox.r and pba.r
- Add-in for Excel (NASA, beta version)
- RAMAS Risk Calc 4.0 (NIH, commercial)
- StatTool (Dan Berleant, freeware)
- PBDemo (NIH, freeware)
- Constructor (Sandia and NIH, freeware)
- Williamson and Downs (1990)

Web presentations and documents

Introduction to probability bounds analysis written for Monte Carlo users
http://www.ramas.com/pbawhite.pdf

Introduction of probability bounds analysis to interval researchers
http://www2.imm.dtu.dk/~km/int-05/Slides/ferson.ppt

Gert de Cooman's gentle introduction to imprecise probabilities
http://maths.dur.ac.uk/~dma31jm/durham-intro.pdf

Fabio's Cozman's introduction to imprecise probabilities
http://www.cs.cmu.edu/~qbayes/Tutorial/quasi-bayesian.html

Notes from a week-long summer school on imprecise probabilities
http://idsia.ch/~zaffalon/events/school2004/school.htm

Introduction to p-boxes and related structures
http://www.sandia.gov/epistemic/Reports/SAND2002-4015.pdf

Handling dependencies in uncertainty modeling
http://www.ramas.com/depend.zip

Introduction to Bayesian and robust Bayesian methods in risk analysis
http://www.ramas.com/bayes.pdf

Statistics for data that may contain interval uncertainty
http://www.ramas.com/intstats.pdf

Topical websites

- Intervals and Probability Distributions
  http://class.ee.iastate.edu/berleant/home/ServeInfo/Interval/intprob.html
- Imprecise Probabilities Project
  http://ippserv.rug.ac.be/home/ipp.html
- Sandia National Laboratory's Epistemic Uncertainty Project
  http://www.sandia.gov/epistemic/
- R software for confidence boxes
  https://sites.google.com/site/confidenceboxes/
- Applied Biomathematics' Risk Calc website
  http://www.ramas.com/riskcalc.htm
- Society for Imprecise Probabilities Theory and Applications
  http://www.sipta.org/

References

Barmish, B.R. and C. M. Lagoa. 1995. The uniform distribution: a rigorous justification for its use in robustness analysis. Technical Report ECE-95-12, Electrical and Computer Engineering, University of Wisconsin, Madison.

Barmish, B.R. and C. M. Lagoa. 1996. The uniformity principle: a new tool for probabilistic robustness analysis. Proceedings of the 30th Annual Conference on Information Science and Systems. Princeton, New Jersey.

Berger, J.O. 1985. *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York.

Ben-Haim, Y. 2001. *Information-gap Decision Theory*, Academic Press, San Diego.

Bukowski, J., L. Korn and D. Wartenberg. 1995. Correlated inputs in quantitative risk assessment: the effects of distributional shape. *Risk Analysis* 15:215-219.

de Cooman, G., T. L. Fine, et al. 2001. *ISIPTA '01. Second International Symposium on Imprecise Probabilities and Their Applications*, Robert Purcell Community Center of Cornell University, Ithaca, NY, USA, Shaker Publishing BV.

Copeland, T.L., A.M. Holbrow, J.M Otani, K.T. Conner and D.J. Paustenbach 1994. Use of probabilistic methods to understand the conservativism in California’s approach to assessing health risks posed by air contaminant. *Journal of the Air and Waste Management Association* 44: 1399-1413.

Couso, I, S. Moral and P. Walley. 1999. Examples of independence for imprecise probabilities. Pages 121-130 in *Proceedings of the First International Symposium on Imprecise Probability and Their Applications*. G. de Cooman, F.F. Cozman, S. Moral and P. Walley (eds.), Imprecise Probabilities Project, Universiteit Gent.

Cox, L.A. Jr., D. Babayev, and W. Huber. 2005. Some limitations of qualitative risk rating systems. *Risk Analysis* 25(3):651-662.

Dempster, A.P. 1967. Upper and lower probabilities induced by a multi-valued mapping. *Annals of Mathematical Statistics* 38: 325–339.

Feller, W. 1948. On the Kolmogorov-Smirnov limit theorems for empirical distributions. *Annals of Mathematical Statistics* 19: 177–189.

Ferson, S., V. Kreinovich, L. Ginzburg, K. Sentz and D.S. Myers. 2002. *Constructing probability boxes and Dempster-Shafer structures*. Sandia National Laboratories, Technical Report SAND2002-4015, Albuquerque, New Mexico. Available at http://www.sandia.gov/epistemic/Reports/SAND2002-4015.pdf.

Ferson, S. 2002. *RAMAS Risk Calc 4.0: Risk Assessment with Uncertain Numbers*. Lewis Publishers, Boca Raton.

Fetz, T. 2001. Sets of joint probability measures generated by weighted marginal focal sets. Pages 171-178 in *Proceedings of the Second International Symposium on Imprecise Probability and Their Applications*. G. de Cooman, T.L. Fine and T. Seidenfeld (eds.), Shaker Publishing, Maastricht.

Frank, M. 1979. On the simultaneous associativity of and $F(x, y)$ and $x+y \square F(x, y)$. *Aequationes Mathematicae* 19: 194-226.

Frank, M.J., Nelsen, R.B., and Schweizer, B. 1987. Best-possible bounds for the distribution of a sum—a problem of Kolmogorov. *Probability Theory and Related Fields* 74, 199-211.

Fréchet, M. 1935. Généralisations du théorème des probabilités totales. *Fundamenta Mathematica* 25: 379-387.

Fréchet, M. 1951. *Sur les tableaux de corrélation dont les marges sont données*. Ann. Univ. Lyon, Sect. A 9: 53-77.

Grosof, B.N. 1986. An inequality paradigm for probabilistic knowledge: the logic of conditional probability intervals. *Uncertainty in Artificial Intelligence*. L.N. Kanal and J.F. Lemmer (eds.), Elsevier Science Publishers, Amsterdam.

Hailperin, T. 1986. *Boole's Logic and Probability*. North-Holland, Amsterdam.

Helsel, D.R. 1990. Less than obvious: statistical treatment of data below the detection limit. *Environmental Science and Technology* 24: 1766–1774.

Insua, D.R. and F. Ruggeri (eds.) 2000. *Robust Bayesian Analysis*. Lecture Notes in Statistics, Volume 152. Springer-Verlag, New York.

Kolmogorov [Kolmogoroff], A. (1941). Confidence limits for an unknown distribution function. *Annals of Mathematical Statistics* 12: 461–463.

Kyburg, H.E., Jr. 1998. "Interval Valued Probabilities," http://ippserv.rug.ac.be/documentation/interval_prob/interval_prob.html, The Imprecise Probabilities Project, edited by G. de Cooman and P. Walley, http://ippserv.rug.ac.be/home/ipp.html.

Lagoa, C.M., and B.R. Barmish. [n.d.] Monte Carlo simulation without probability distributions: a tutorial. [unpublished manuscript].

Lagoa, C.M., P.S. Shcherbakov and B.R. Barmish. 1998. Probabilistic enhancement of classical robustness margins: the unirectangularity concept. *Systems and Control Letters* 35: 31-43.

Loui, R.P. 1986. Interval based decisions for reasoning systems. *Uncertainty in Artificial Intelligence*. L.N. Kanal and J.F. Lemmer (eds.), Elsevier Science Publishers, Amsterdam.

Mishra, S. 2000. Uncertainty propagation using the point estimate method. *Calibrated Reliability in Groundwater Modeling*, IAHS.

Moore, R.E. 1966. *Interval Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey.

Moore, R. 1979. *Methods and Applications of Interval Analysi*s. SIAM, Philadelphia.

Nelsen, R.B. 1999. *An Introduction to Copulas*. Lecture Notes in Statistics 139, Springer-Verlag, New York.

Rosenblueth, E. 1975. Point estimates for probability moments. *Proceedings of the National Academy of Sciences USA* 72: 3812-3814.

Rowe, N.C. 1988. Absolute bounds on the mean and standard deviation of transformed data for constant-sign-derivative transformations. *SIAM Journal of Scientific Statistical Computing* 9: 1098–1113.

Rustichini, A. 2005. *Science* 310: 1624.

Shafer, G. 1976. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ.

Shlyakhter A. 1994. Improved framework for uncertainty analysis: accounting for unsuspected errors. *Risk Analysis* 14(4):441-447.

Smith, A.E., P.B. Ryan and J.S. Evans. 1992. The effect of neglecting correlations when propagating uncertainty and estimating the population distribution of risk. *Risk Analysis* 12:467-474.

Tessem, B. 1992. Interval probability propagation. *International Journal of Approximate Reasoning* 7: 95-120.

Vesely, W.E., F.F. Goldberg, N.H. Roberts, D.F. Haasl. 1981. *Fault Tree Handbook*. Nuclear Regulatory Commission, Washington, DC.

Walley, P. 1991. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London.

Walley, P. 1996. Inferences from multinomial data: learning about a bag of marbles. *Journal of the Royal Statistical Society* 58: 3-37.

Walley, P. 2000. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning* 24: 125-148.

Walley, P. and T.L. Fine 1982. Towards a frequentist theory of upper and lower probability. *Annals of Statistics* 10: 741-761.

Whitt, W. 1976. Bivariate distributions with given marginals. *The Annals of Statistics* 4:1280-1289.

Williamson, R.C. and Downs, T. 1990. Probabilistic arithmetic I: Numerical methods for calculating convolutions and dependency bounds. *International Journal of Approximate Reasoning* 4 89-158.

Wise. B.P. and M. Henrion 1986. A framework for comparing uncertain inference systems to probability. *Uncertainty in Artificial Intelligence*. L.N. Kanal and J.F. Lemmer (eds.), Elsevier Science Publishers, Amsterdam.

Yager, R.R. 1986. Arithmetic and other operations on Dempster-Shafer structures. *International Journal of Man-machine Studies* 25: 357-366.