Real-world case studies using Bayesian analysis with @RISK, and how to use Bayesian Analysis in your models

Dr. Francisco J Zagmutt
Managing Partner
EpiX Analytics

Sound answers for complex decisions

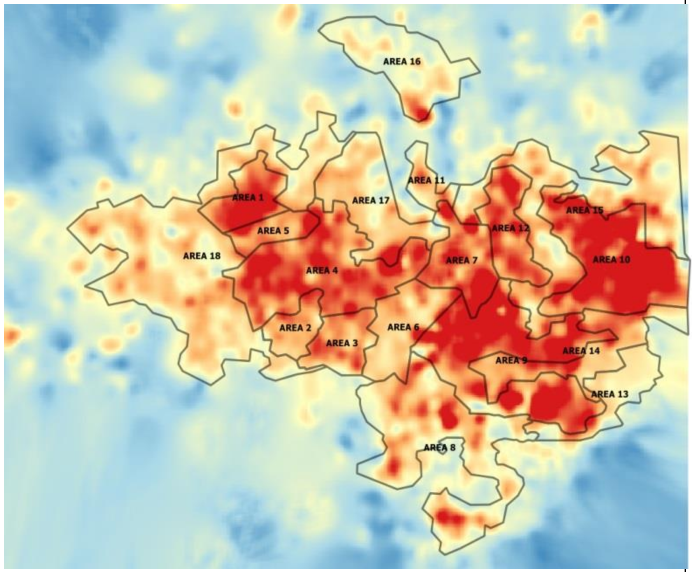
www.epixanalytics.com

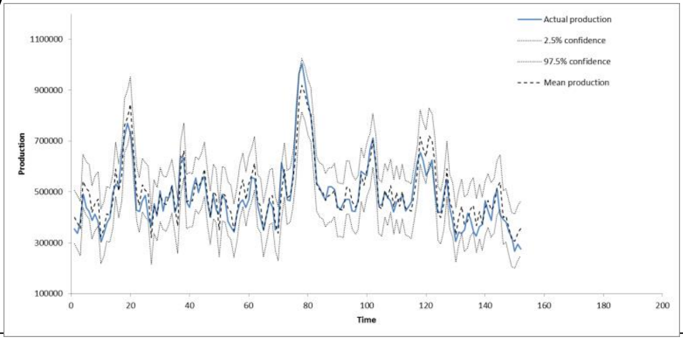
Who is this guy and his company?

1. Managing partner at EpiX Analytics - specialized risk analytics and decision-modeling company

2. *Focus*: Quantitative risk analysis & modeling *to* improve decision-making

3. Experience in a wide range of industries:
   - Pharmaceuticals
   - Mining
   - Manufacturing
   - Transportation
   - Insurance
   - Outcomes research / pharmacoeconomics
   - Financial industry
   - Health / Food safety
   - Energy, oil & gas
   - Many others….





EPIX ANALYTICS

Some of the institutions we have helped

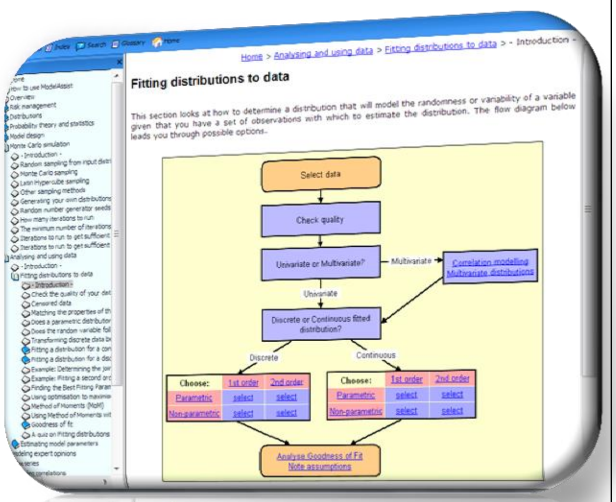
JOHN DEERE
KPMG
ECOLAB
Unilever
citi
The Coca-Cola Company
AON
Palamon Capital Partners
CORNING
Johnson & Johnson
ConocoPhillips
xerox
U.S. Food and Drug Administration
Protecting and Promoting Your Health
FREEPORT-MCMORAN COPPER & GOLD
Pfizer
gsk GlaxoSmithKline
ABInBev
ABN·AMRO
Agilent Technologies
Lilly
EDRINGTON
NEWMONT The Gold Company
Juniper NETWORKS
TEVA
Spectra Energy
USDA United States Department of Agriculture
FEDERAL COMMUNICATIONS COMMISSION U S
EMERSON Network Power
Duke Energy

ModelAssist for @RISK

<u>Free</u> simulation and @RISK training and reference tool:

http://www.epixanalytics.com/ModelAssist.html

Page numbers are *Mxxxx*. For example, *M0407* is “*Selecting the appropriate distributions for your model*”



Keep an eye on these during the seminar!

Introduction

Bayesian analysis - from a niche academic technique to the forefront of risk analysis and predictive analytics

But implementation can be intimidating

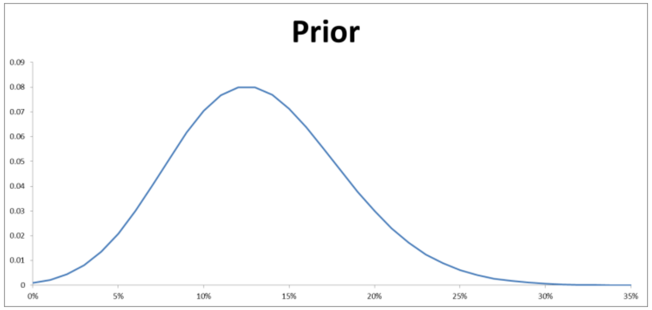
> Literature mostly using statistical packages or specialized tools rather than general modeling tools like @RISK

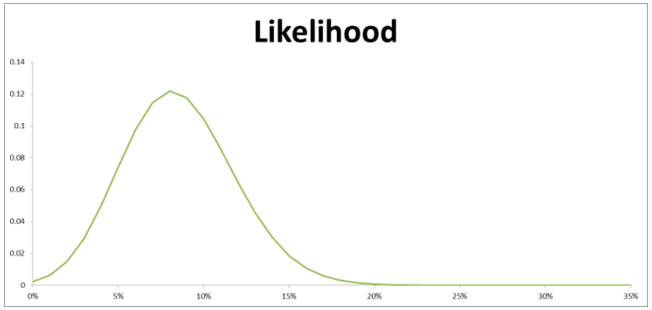
Goal - gentle introduction to Bayesian analysis, and how to use it in @RISK based on real-life case studies and example models

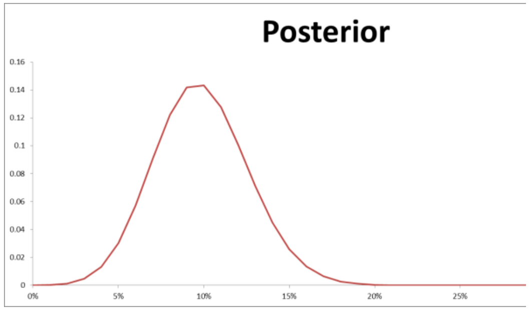
Emphasis on incorporating Bayesian methods within a risk model without affecting its general structure

The principle of bayesian analysis

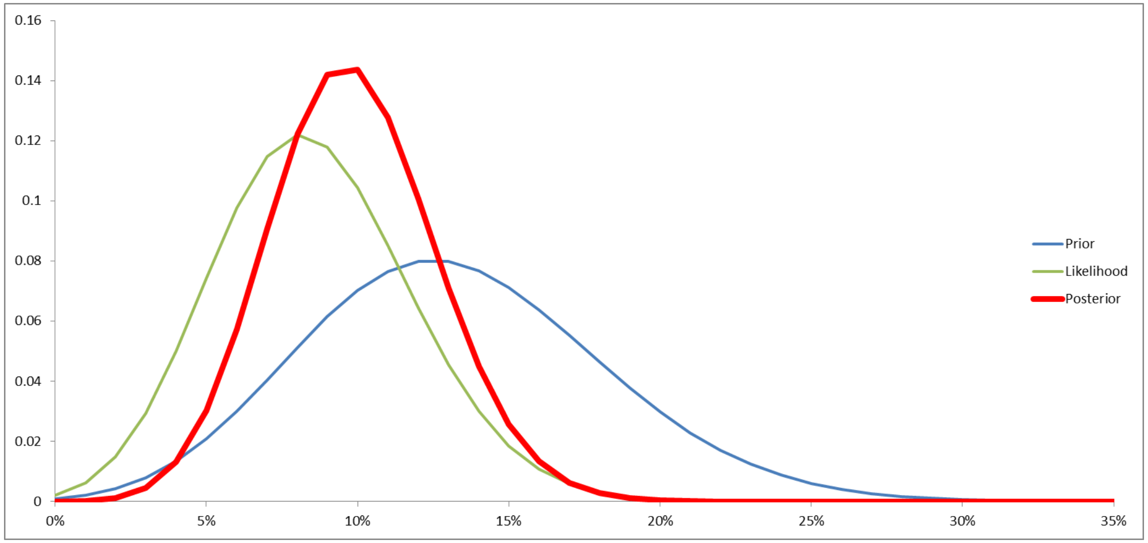
Parameters and distributions are estimated using:







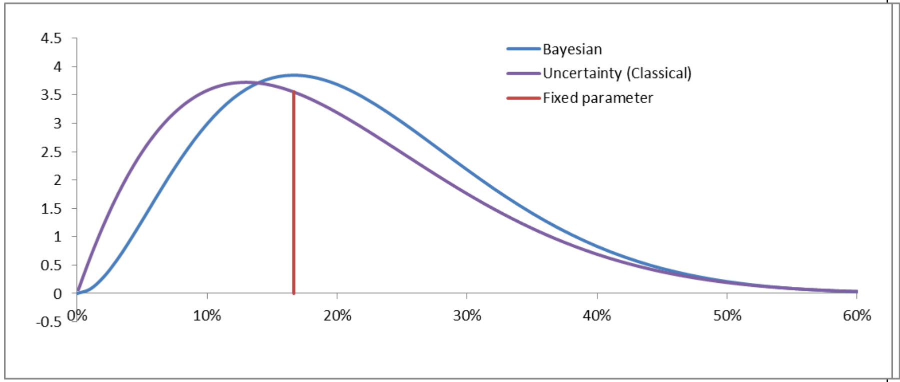
Prior knowledge + Current data = Current knowledge



Difference with "classical statistics?

Classical, Fisherian, or frequentist statistics key foundations:

- Parameters are **fixed**, but may be unknown (the infamous "error") – e.g. $SEM=\sigma/\sqrt{n}$
- Can only be estimated using **current data**



Bayesian statistics:

- Parameters are **distributions**
- Estimation is based on current belief of parameters, based on **past** belief and **current** data

Right, great subject for a dinner party…but how does Bayesian help me?

- Easy way to combine multiple sources of information
- More flexible uncertainty modeling, particularly under scarce data
- Results essentially equal with enough data

Bayesian analysis

Not only used by Google….

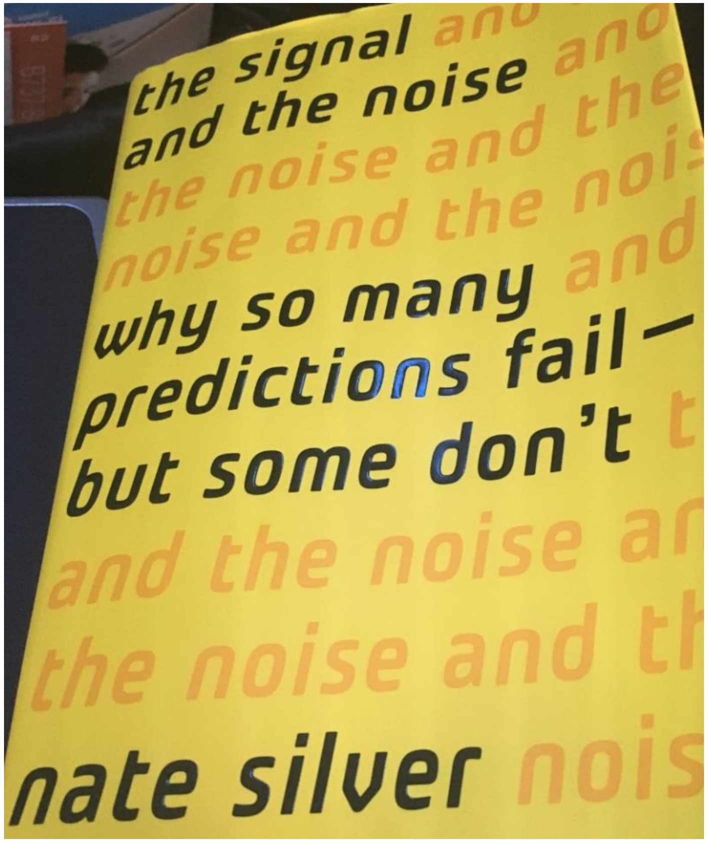
Central premise:

Most forecasts are wrong because they only take into account historical data/patterns

How do we reconcile the need to use the past as a guide with our recognition that the future may be different?

Main solution:

Think probabilistically and Bayesian methods



Applying Bayes' theorem

$$P(B|A) = \frac{P(B)P(A|B)}{P(A)}$$

$$f(\theta|X) = \frac{\pi(\theta).l(X|\theta)}{\int \pi(\theta).l(X|\theta).d\theta}$$

$\pi(\theta)$ - the "prior distribution", = distribution of prior knowledge about θ.

$l(X|\theta)$ - the "likelihood function", = calculated probability of observing the data ***X*** for a given value of θ.

$f(\theta|X)$ - the "posterior distribution", = distribution of current knowledge of θ after we have observed the data ***X*** and given knowledge of the value of θ before ***X*** was observed.

The denominator above is just a normalizer so this simplifies to:

$$f(\theta|X) \propto \pi(\theta).l(X|\theta)$$

Prior distribution types

Conjugate prior - a prior that is the mathematical conjugate of the likelihood function (i.e. it has the same basic form). Simplifies mathematics a lot

Non-informative (uninformed) prior - contains "no information". A statement of ignorance prior to the observations. **Vaguely informative** more appropriate name since some assumptions always built into prior.

Elicited prior - a distribution determined by expert opinion

Conjugate priors (M0101)

Shortcut to calculate a posterior in one simple pass

Works when posterior and prior distribution have the **same functional form**

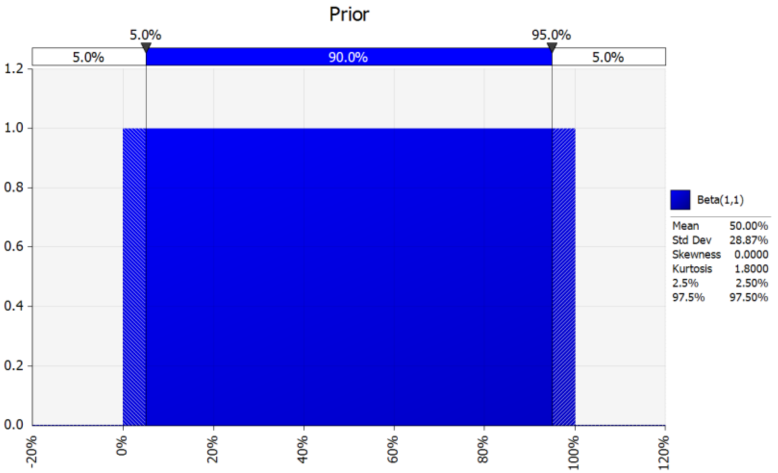
e.g. **Beta** is conjugate to **Binomial,** so when we observe s successes from n trials we can model the probability of success P using **Riskbeta(s+α, n-s+β)**

If we don't have prior information on P, $\alpha = \beta = 1$ (see next slide)

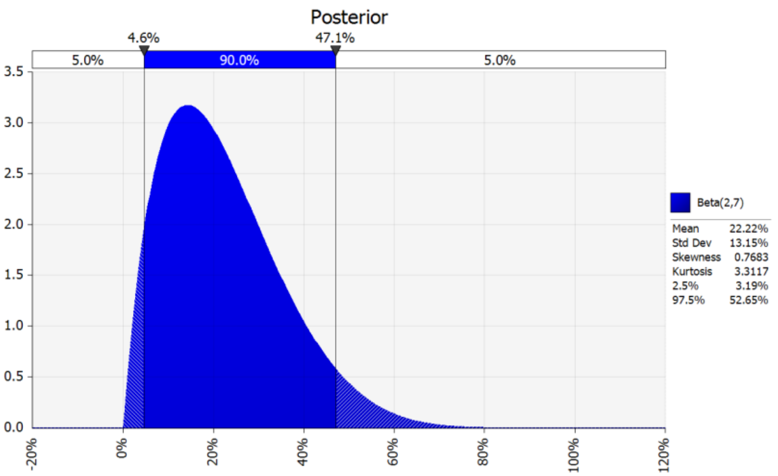
Example - modeling **uncertainty in proportions** (M0171)

Contractor has gone over budget in 1 out of the 7 projects we have used them. Probability of going over budget in next project?

(**Non-informative**) Prior = RiskBeta(1, 1)



Posterior=RiskBeta(s+1,n-s+1) => RiskBeta(1 +1, 7 - 1 +1)



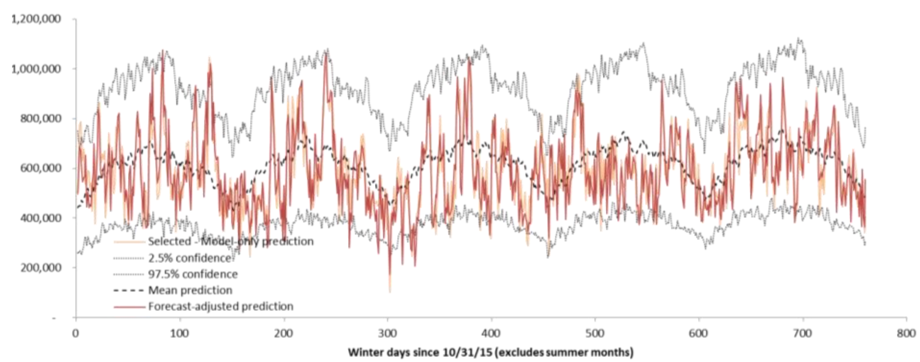
ContractorProbabilityBudgetOverrun.xlsx

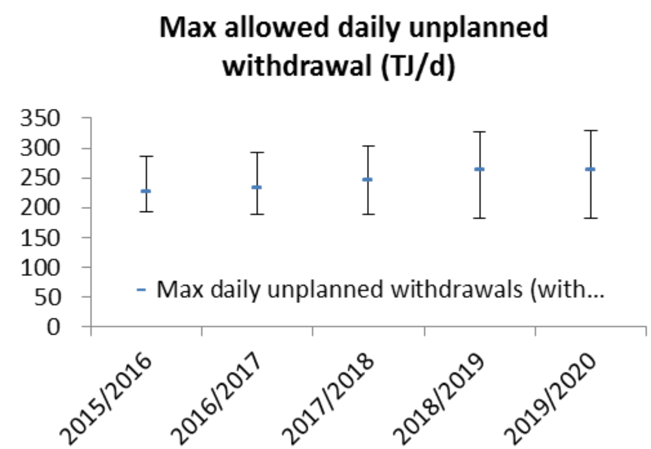
Conjugate priors on a real application: Case 1 - Bayes keeps you warm

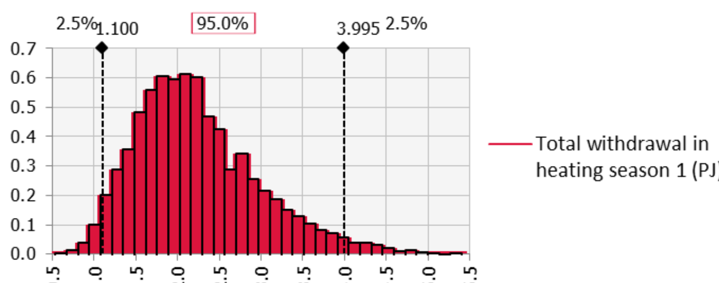
Regulated utility needed to forecast critical **gas** storage capacity for mid-term planning

Getting it wrong not only very costly but potentially life threatening

EpiX developed sophisticated @RISK model that incorporated historical demand and weather data, engineering inputs, and results from a system equilibrium model to forecast daily demand/withdrawals for several years

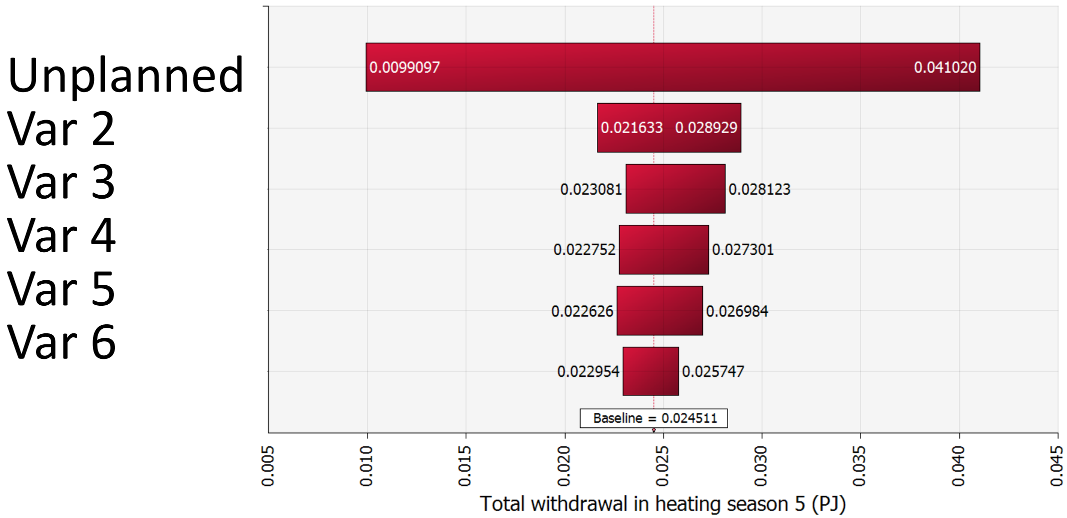






Capacity uncertainty significantly driven by unplanned withdrawals…

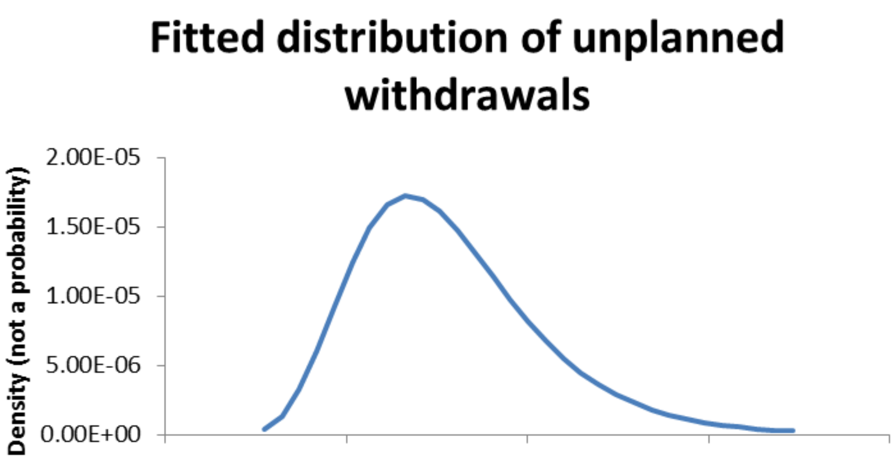
But only few unplanned withdrawals stored in current database

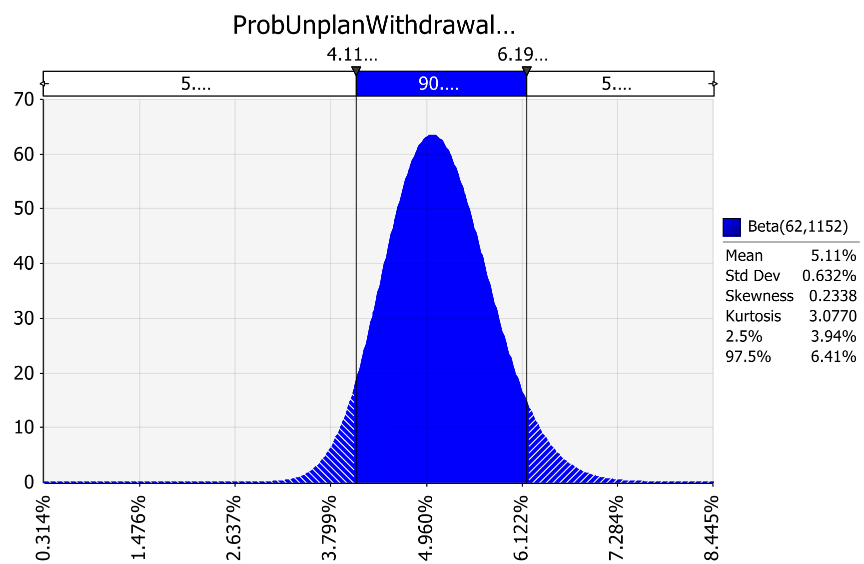


[BayesKeepsYouWarm.xlsx]

Approach

- Withdrawal amount: Fitted to historical data, with uncertainty using ***classical statistics***
- Daily probability of withdrawal (P_w): using Beta(s+1, n-s+1)
  - s = # of unplanned withdrawals observed
  - n= total days in dataset
- Then for each day RiskBernoulli(P_w)*RiskLognormal(μ,σ)





Case study 2 – conjugate priors and equipment failure

Client produces and services critical (and expensive) equipment

Significant profits from service contracts

EpiX created models to **forecast** services and headcounts, and **optimize parts inventory** while taking into account key business drivers, including **equipment in field**

Subset of models required answering following questions:
- Do some machines fail more often?
- Can we rank them, with uncertainty?
- How confident are we in the results?

Example data and approach

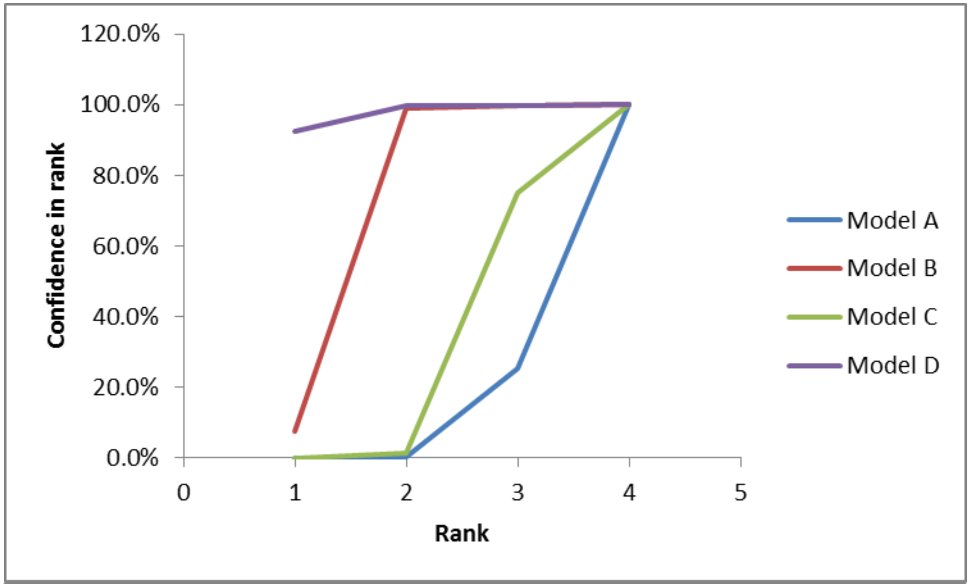
Type	Total machine/days in service	Failures	Mean rate/year
Model A	38,325	134	1.28
Model B	10,950	22	0.73
Model C	78,475	255	1.19
Model D	3,066	3	0.36

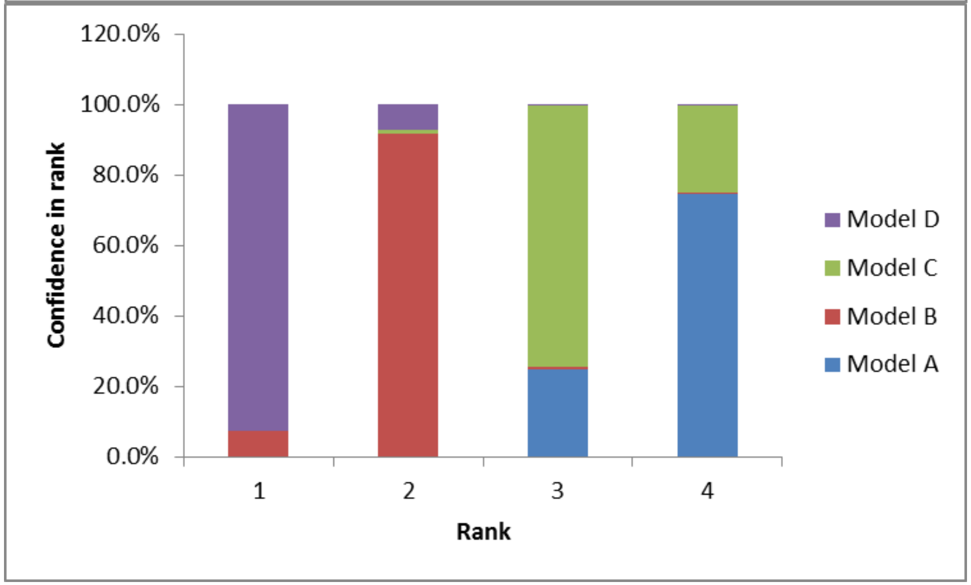
We can use the following conjugate prior to model failure rate for each model: ***Gamma(observed events, 1/exposure)***, thus

=RiskGamma(Failures, 1/machine days in service)

Then compare rates using simulation – calculate confidence in lowest rate, and also on ranks.

MachineFailures.xlsx

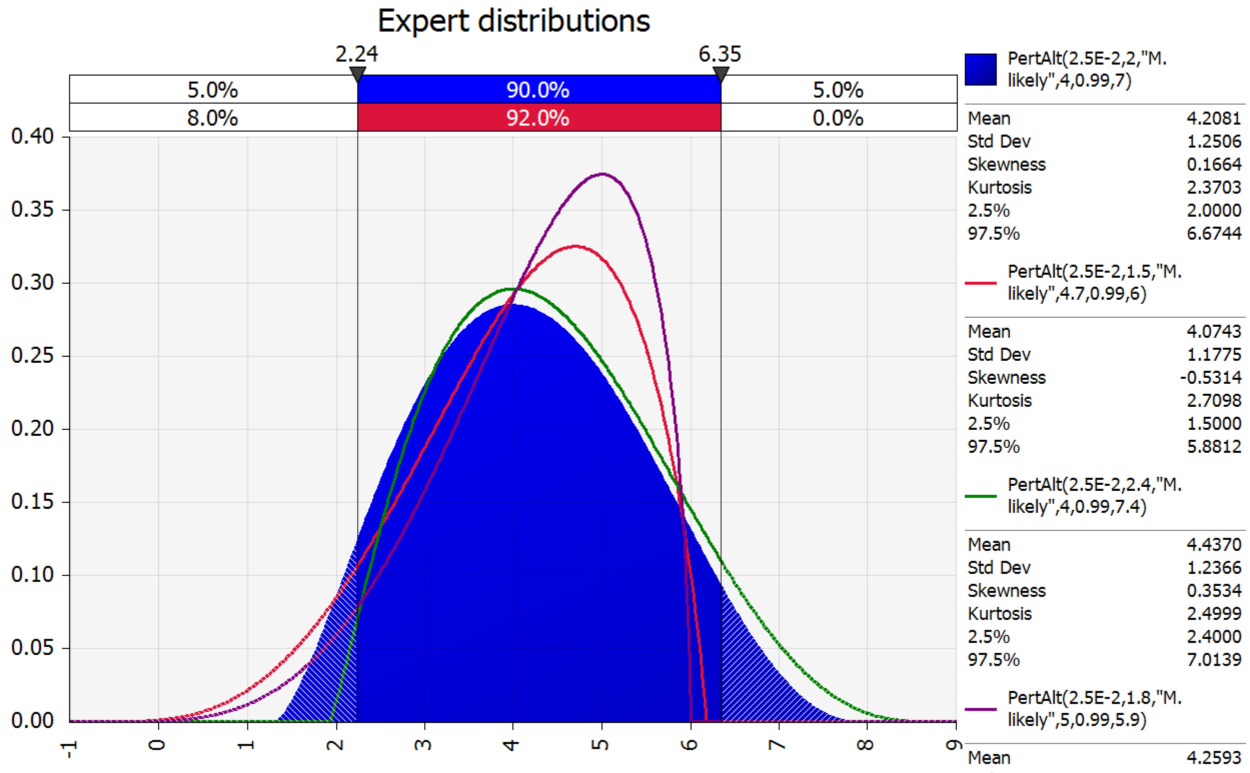




Case 3 - Combining expert opinions in project RA (M0229)

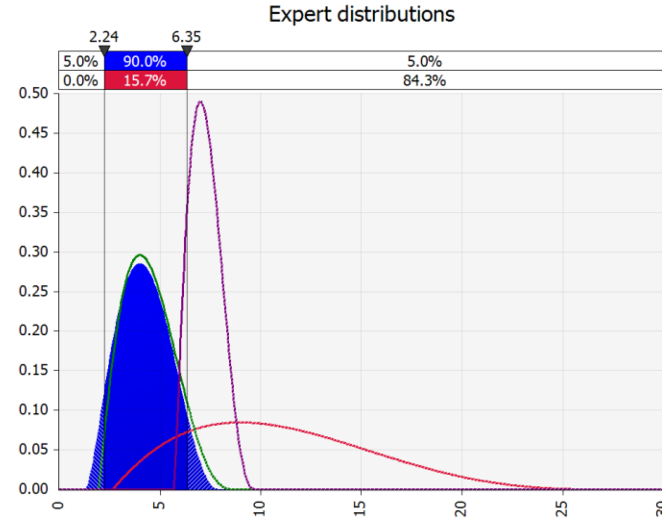
Bayesian analysis can be used to combine multiple information, including expert opinion

If expert's opinion agree, standard combination of all experts might makes sense (e.g. Meta-analysis)



But what if they disagree? Bayesian model averaging (BMA)

Cost (M) to complete project step				
	2.50%	**M. likely**	**99%**	**Expertise weights**
Expert 1	2	4	7	1
Expert 2	4	9	22	3
Expert 3	2.4	4	7.4	1
Expert 4	6	7	9	2



Idea is to create a mixture of expert opinion, weighted by their *likelihood*

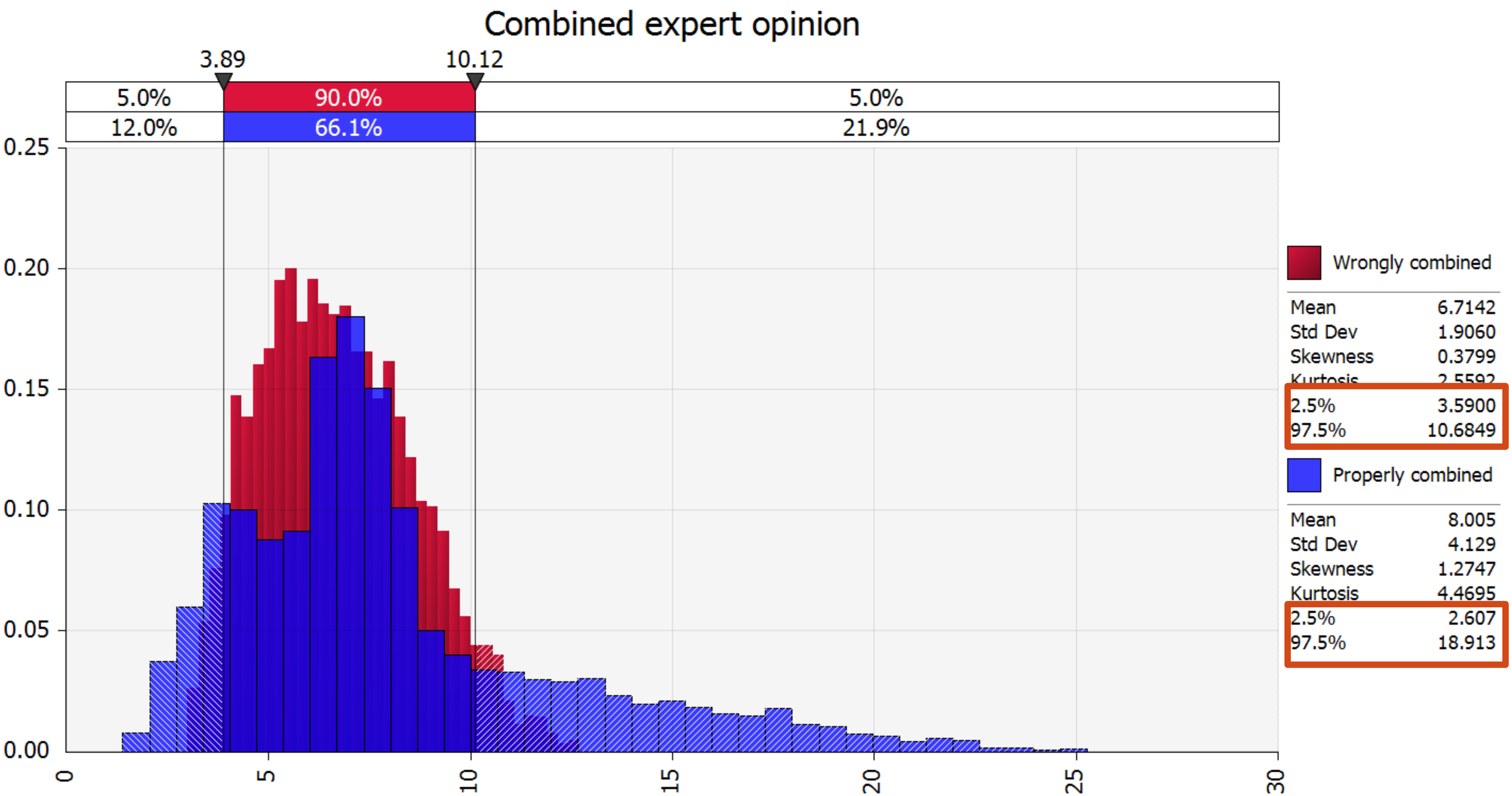
As no direct *likelihood* from expert opinion, we can instead use subjective **weights** based on expertise assessment

We then use the **RiskDiscrete({Expert distributions}, {Weights})** to combine them

What happens if we just take the mean of 2.5, ML and 99% for all experts, then use in RiskPertAlt?

Comparing methods

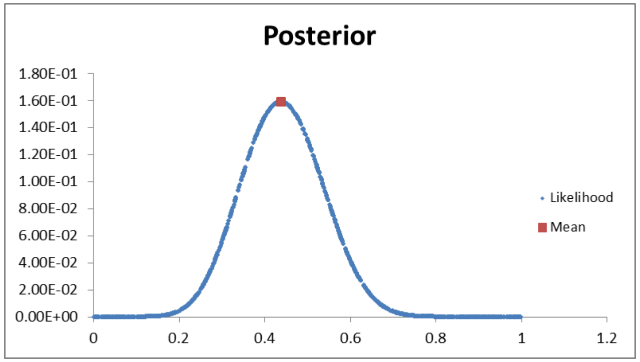
[Experts_Disagree.xlsx]



Other methods (M0052)

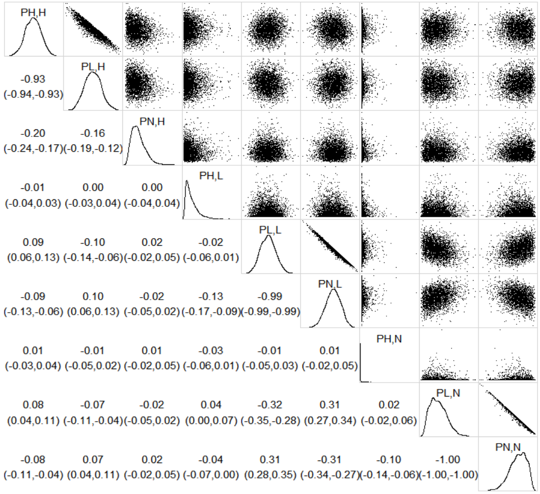
Single parameters

- Conjugate – already reviewed
- Construction of posterior



Single or **multiple** parameters

- Bayesian Monte Carlo
- Approximate Bayesian Computation (ABC)
- Markov Chain Monte Carlo (MCMC)
- Integrated Nested Laplace Approximation (INLA)



Comparison of methods BinomialBayesianMC_@RISK.xlsx

Conclusions

Only method that allows us to incorporate past, data, and expert opinion in our estimate -> more accurate predictions

Easy to implement in @RISK without disrupting your model structure

A model can include both Bayesian and classical methods

Limitation: when multiple parameters are correlated, harder to implement in @RISK, but possible. See M0052

Time to practice with your own models!

Thanks for your time!

Please contact me with any questions, or for **a free copy of the models and/or this talk**

Francisco J Zagmutt, DVM, MPVM, PhD
Managing Partner
EpiX Analytics
fzagmutt@epixanalytics.com