EXPECTED PROBABILITY, MEAN HAZARD, POSTERIOR PREDICTIVE

DLS-216, Module 24

US Army Corps of Engineers®
Dam and Levee Safety Program

U.S.ARMY®

March 2026/Version 1

GAVIN'S POINT DAM, SD (SOURCE: USACE)

Learning Objectives

- Define expected probability, mean hazard, and posterior predictive curves
- Describe how expected value is calculated
- Demonstrate how expected probability (posterior predictive) curves is calculated for Bayesian analysis
- Explain why we use the mean to inform risk

Expected Value

Anticipated return from a set of numbers and their corresponding probabilities

Expected value $E(x) = \sum x * P(x)$

Normal 6-sided die $E(x) = 1/6 * (1+2+3+4+5+6)$

$E(x) = 3.5$

X	P(x)
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

Weighted die twice as likely to roll a 1, 2, or 3 than a 4, 5, or 6

$E(x) = 2/9 * (1+2+3) + 1/9 * (4+5+6)$

$E(x) = 3$

X	P(x)
1	2/9
2	2/9
3	2/9
4	1/9
5	1/9
6	1/9

Expected Value Example

Calculate the expected value of the following dataset

$$E(x) = \sum x * P(x)$$

X	P(X)
77	0.1
80	0.2
91	0.3
99	0.4

Expected Value Example

$$E(x) = \sum x * P(x)$$

x	P(x)	x * P(x)
77	0.1	7.7
80	0.2	16
91	0.3	27.3
99	0.4	39.6
		90.6

Expected Value Example II

Calculate the expected value of the following dataset

$$E(x) = \sum x * P(x)$$

x	P(x)
0.01	0.25
0.001	0.25
0.0001	0.25
0.00001	0.25

Expected Value Example II

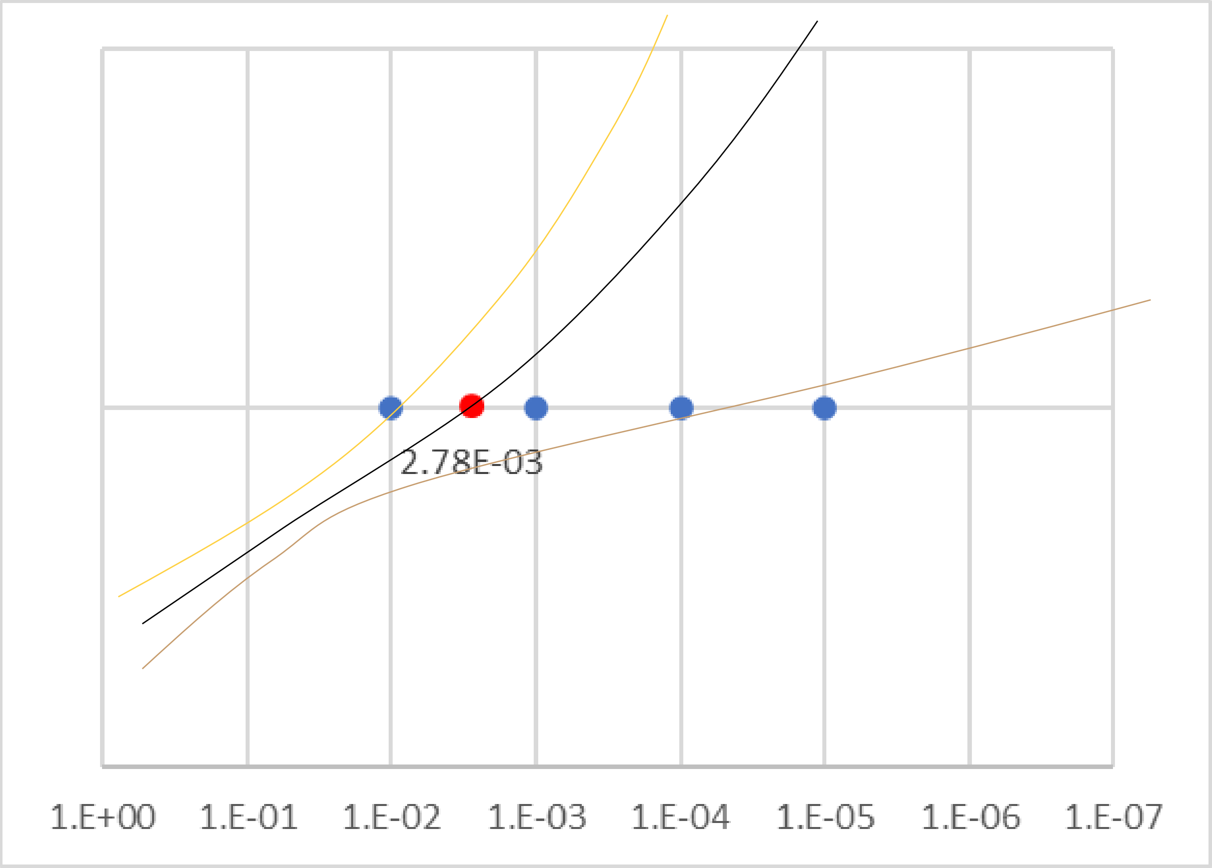
$$E(x) = \sum x * P(x)$$

x	P(x)	x * P(x)
0.01	0.25	2.50E-03
0.001	0.25	2.50E-04
0.0001	0.25	2.50E-05
0.00001	0.25	2.50E-06
		2.78E-03

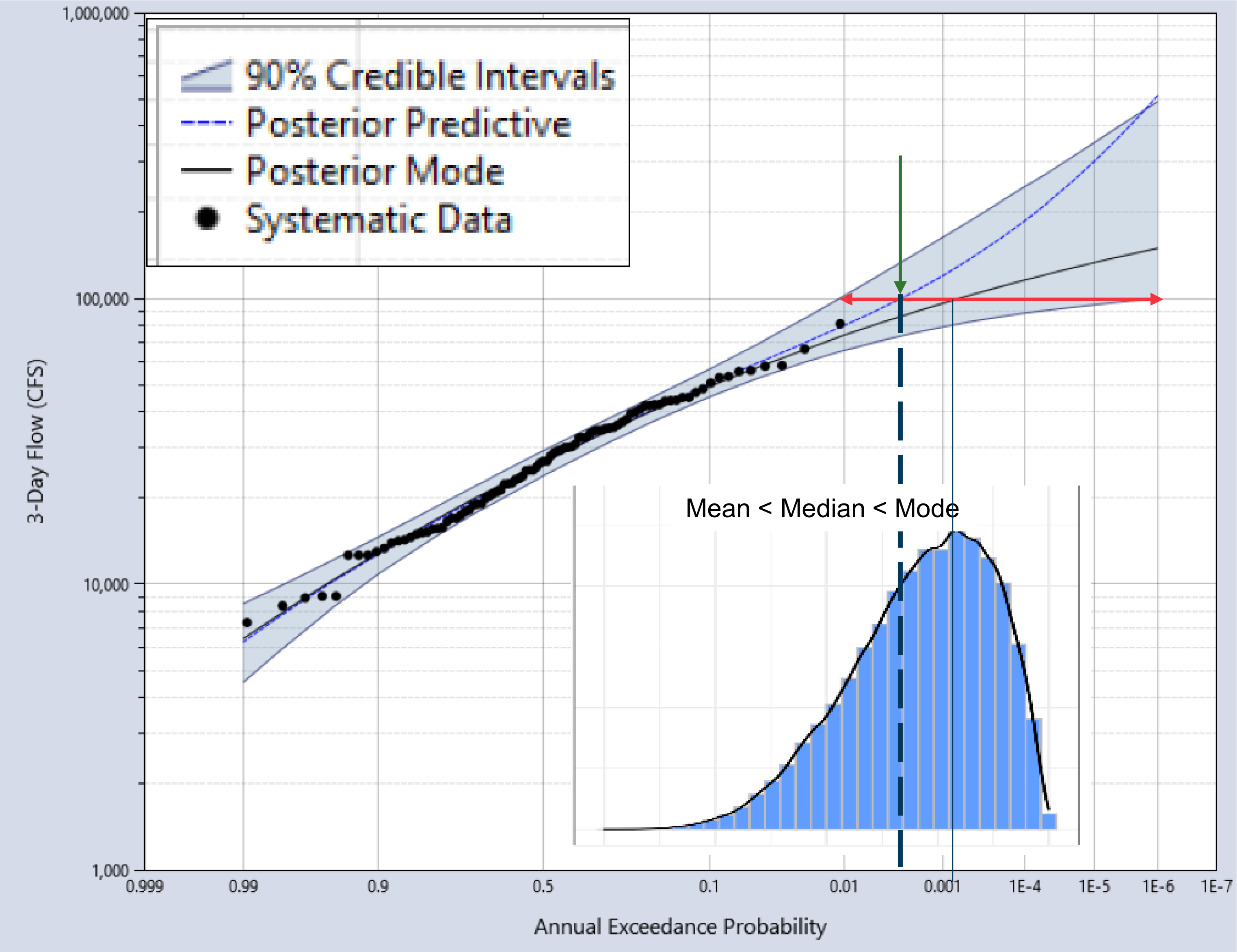
Expected Value Example II

$$E(x) = \sum x * P(x)$$

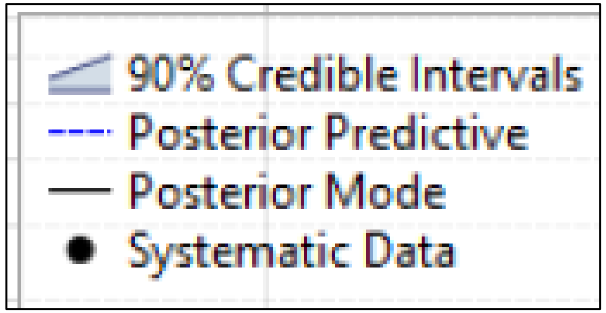
x	P(x)	x * P(x)
0.01	0.25	2.50E-03
0.001	0.25	2.50E-04
0.0001	0.25	2.50E-05
0.00001	0.25	2.50E-06
		2.78E-03



Expected Probability



Posterior Distributions



Prior distribution + new information = posterior distribution

Prior distribution (default parameter distribution, regional skew, precip frequency)

+

New information (systematic, historic events, record extension, paleoflood)

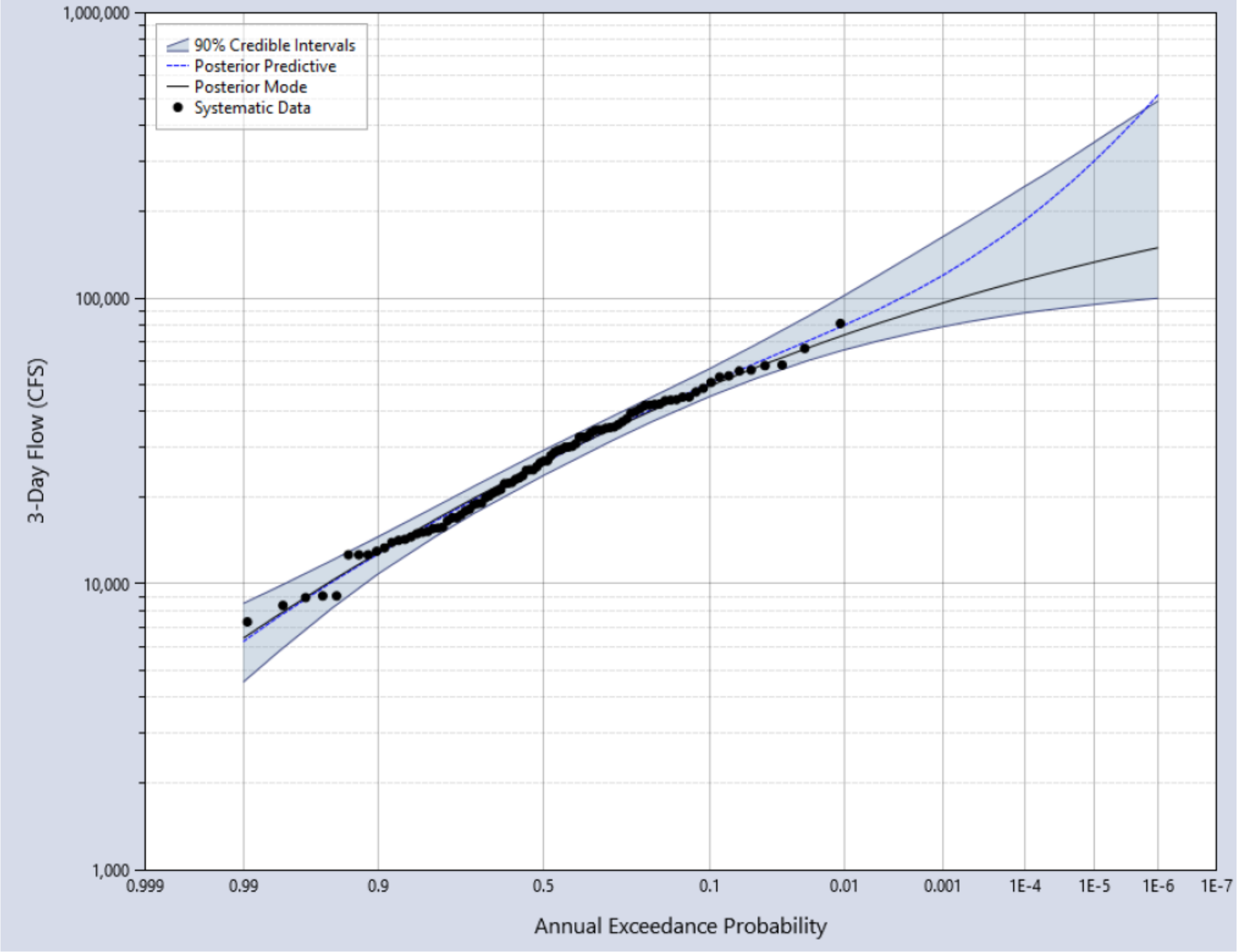
=

Posterior distribution (volume frequency curve)

Not All Errors Created Equal

Why predictive value?

- Avoid exceeding your target value more than intended
- As sample increases, ERL increases, the predictive value converges with observed data



Calculating the Predictive Value

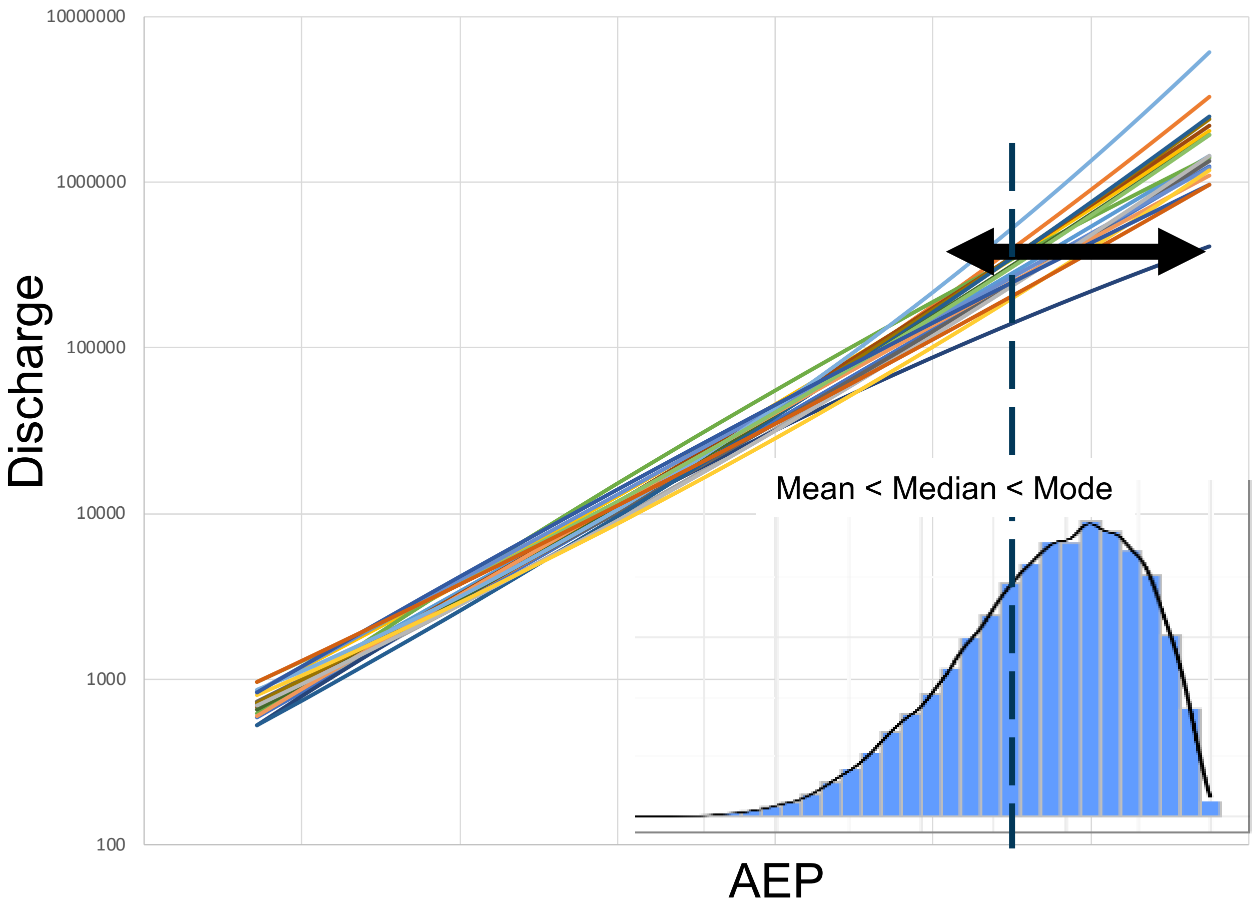
- For a given discharge:

$$E(x) = \sum x * P(x)$$

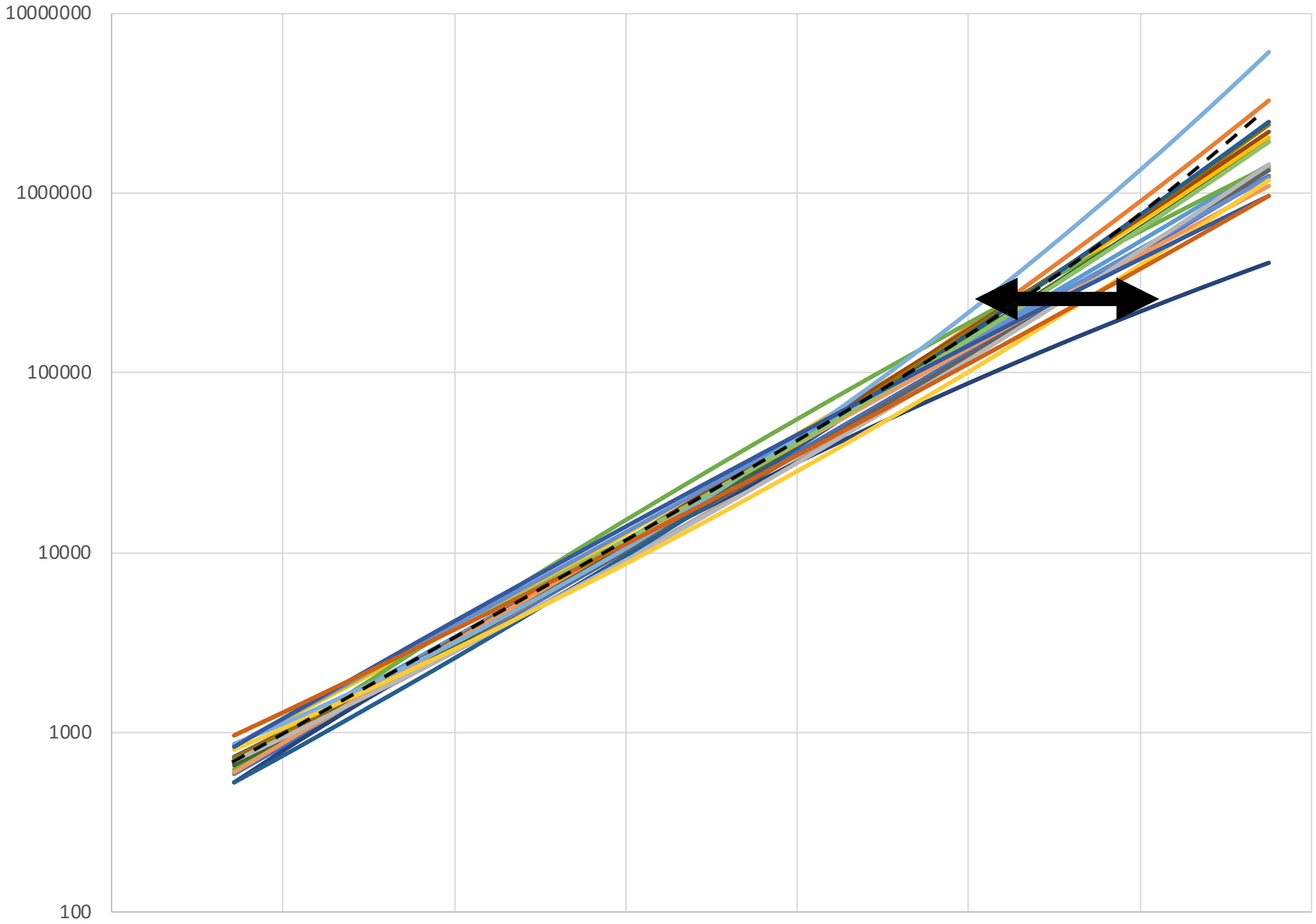
AEP 1/n

- all distributions are equally likely
- expected value simplifies to the arithmetic mean

$$E(x) = \sum x / n$$

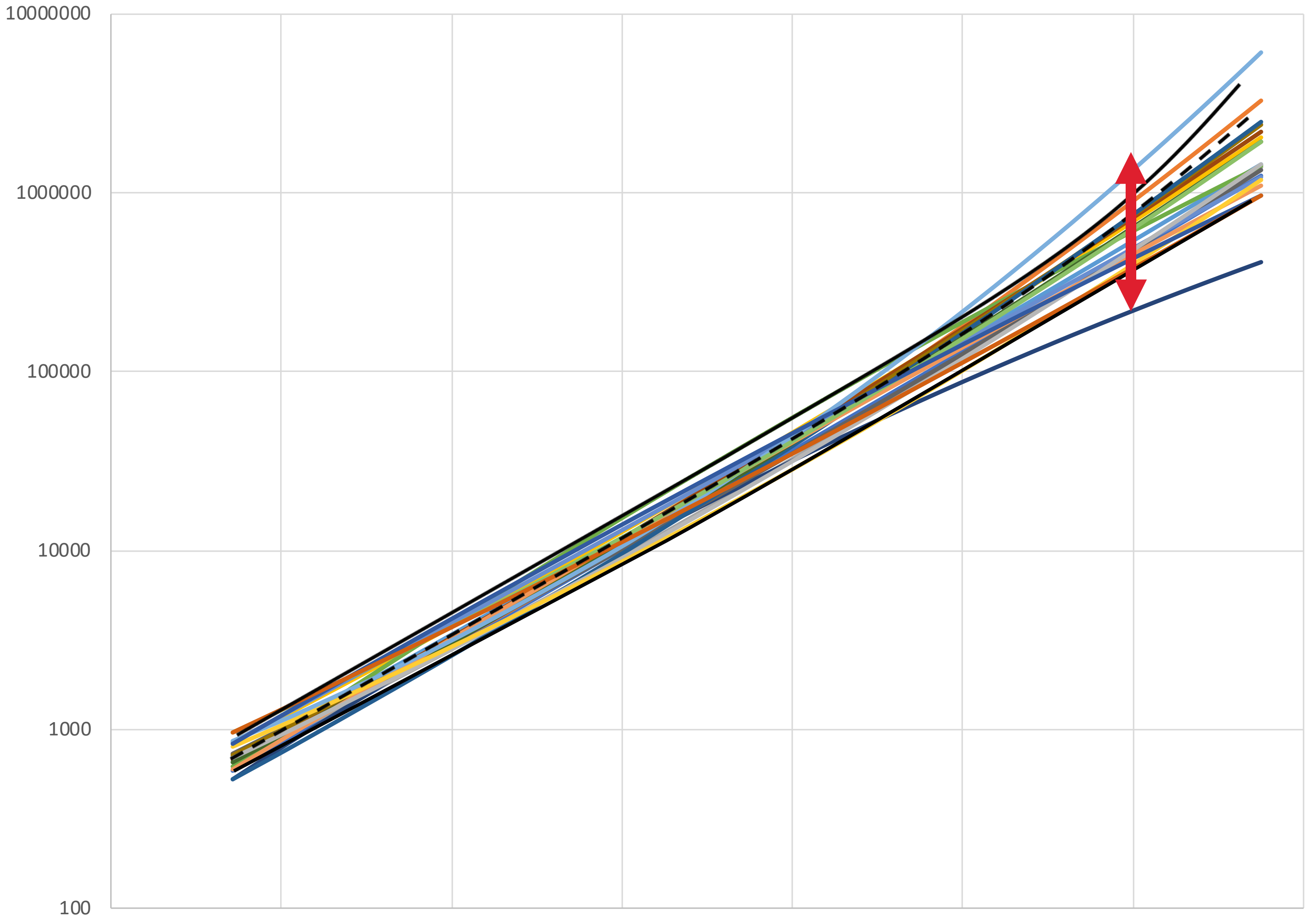


Calculating the Predictive Value



x	P(x)
6.1E-5	0.05
7.3E-5	0.05
8.2E-5	0.05
6.9E-5	0.05
2.7E-5	0.05
1.3E-5	0.05
8.9E-6	0.05
7.1E-5	0.05
…	…
5.0E-5	0.05

Calculating the Predictive Value



Discharge	Percentile
2.2E6	1.00
1.0E6	**0.95**
9.8E5	0.90
…	…
5.2E5	0.30
4.9E5	0.25
4.6E5	0.20
4.2E5	0.15
4.1E5	0.10
2.3E5	**0.05**

Why Do We Use Expected Probability?

- The expected frequency curve “can serve as a basis for computing the expected return on investment” (Beard, 1960)
- The expected flow-frequency curve is the optimal estimator in the context of flood hydrology and should be used to inform decisions (USACE, 1994)
- If the top of levee is equal to the computed (Median or Mode) 100-year flood level, then the levee will overtop MORE frequently (on average) than once every 100 years. (not good)
- If the top of levee is equal to the expected (mean) 100-year flood level, then the levee will overtop (on average) once every 100 years. (Good!)

Summary

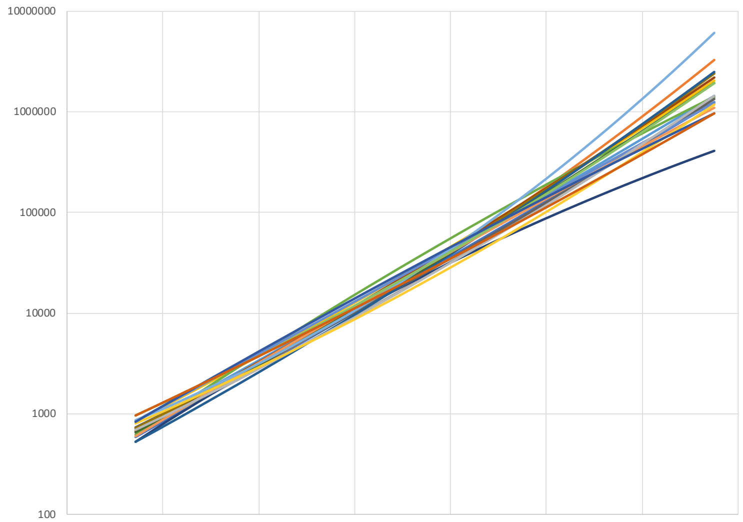
- Posterior predictive calculated from equally probable distributions
- Since all distributions have the same probability, the expected value is the mean
- Credible intervals are calculated across the distributions as well.
- We use expected probability to inform risk

Expected Probability
Mean Hazard Curve
Posterior Predictive

Predictive

5th 95th

Rank order percentile



Learning Objectives Recap

- Define expected probability, mean hazard, and posterior predictive curves
- Describe how expected value is calculated
- Demonstrate how expected probability (posterior predictive) curves is calculated for Bayesian analysis
- Explain why we use the mean to inform risk

Questions