

# Society of Petroleum Evaluation Engineers

## Monograph 3-

### Chapter 2

## **“Statistics-A Brief Lesson”**

# Sum of Independent Distributions

Die One

Likelihood

1	2	3	4	5	6

Die Two

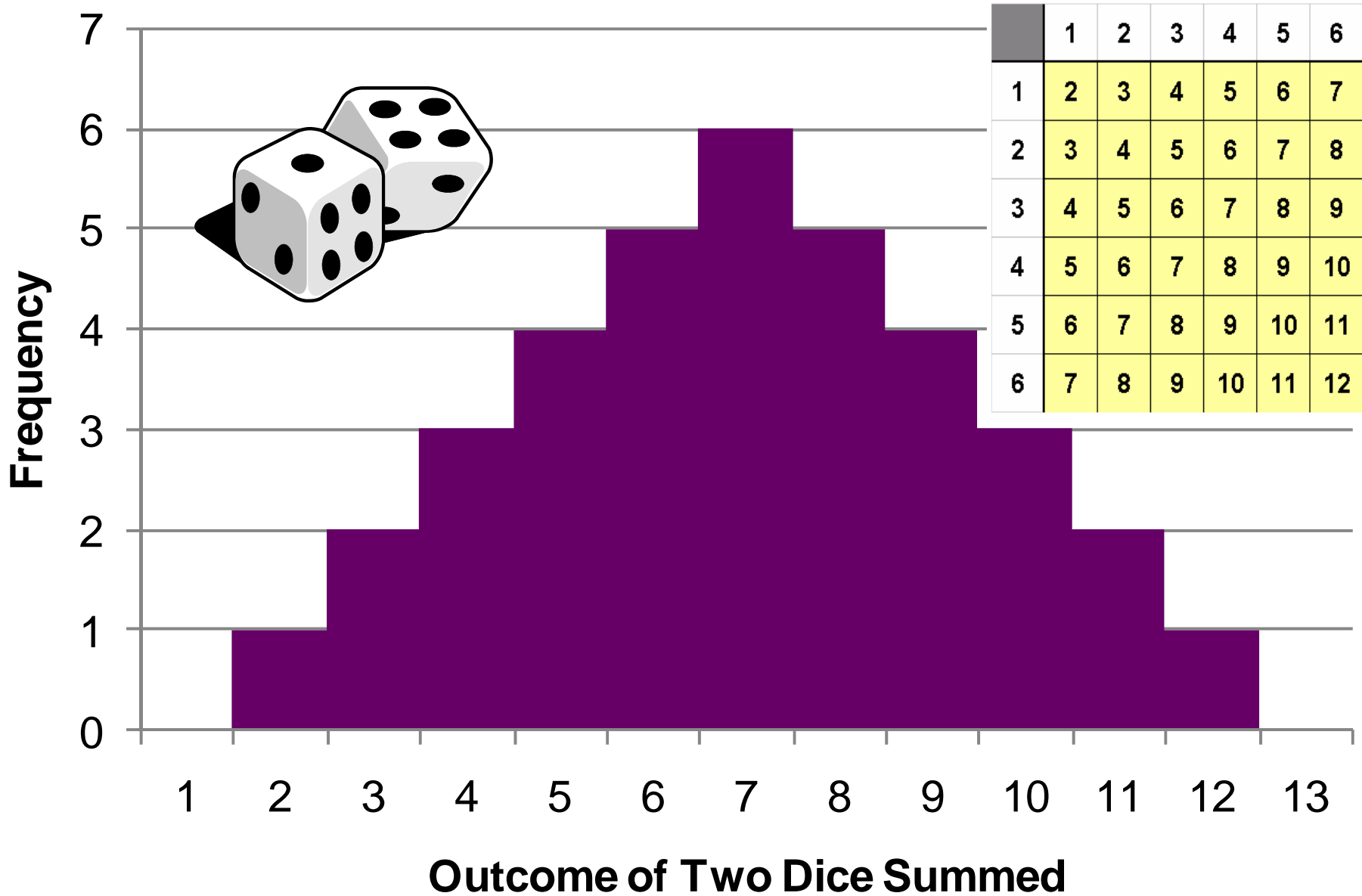
Likelihood

1	2	3	4	5	6

Outcome



# Sum of Independent Distributions



# Product of Independent Distributions

Die One

Likelihood

1	2	3	4	5	6

Die Two

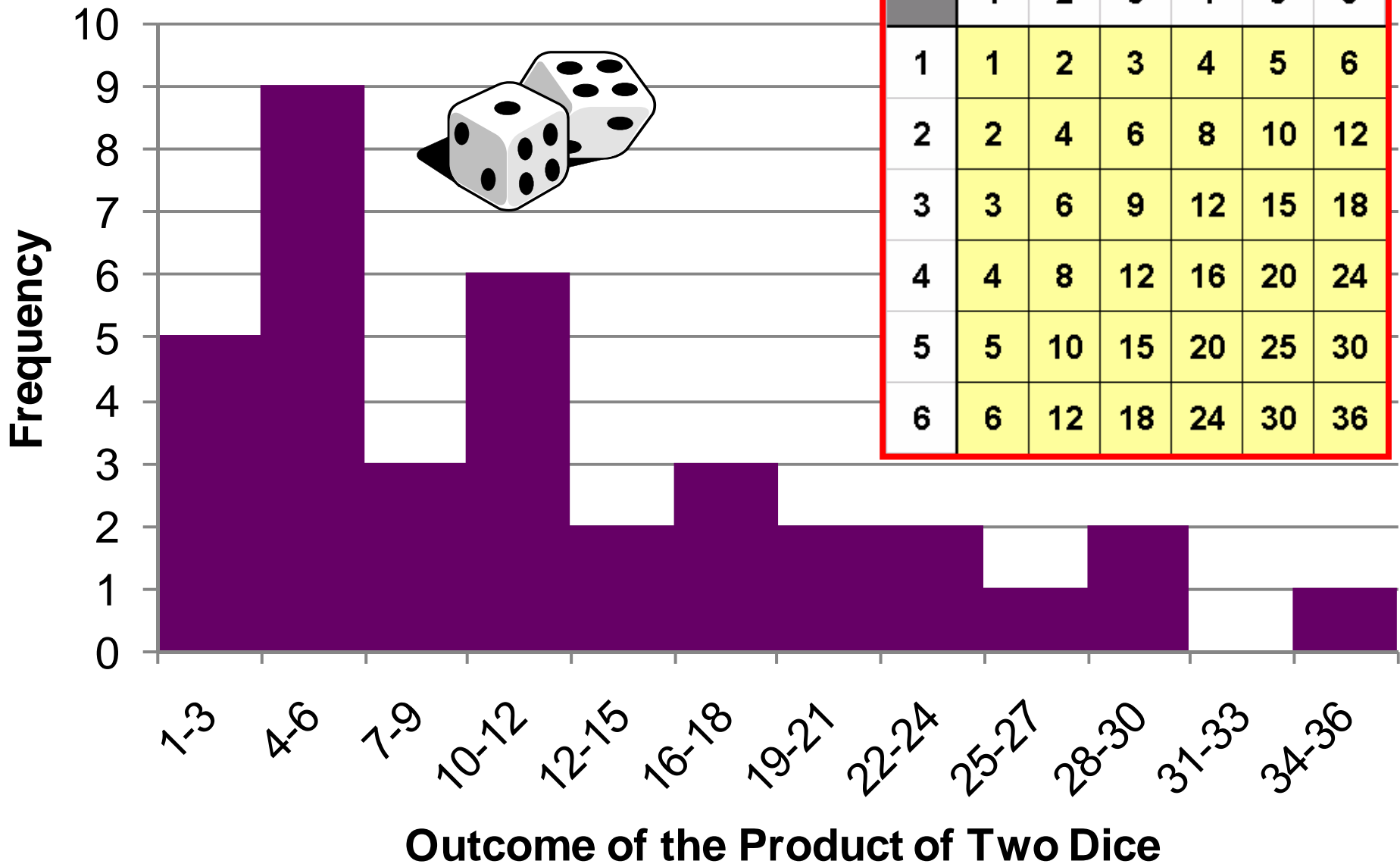
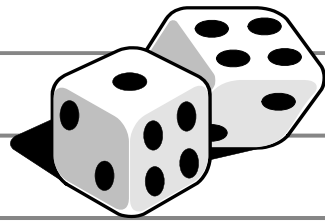
Likelihood

1	2	3	4	5	6

Outcome

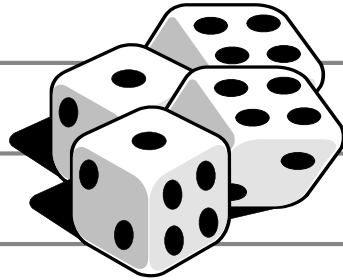


# Two Dice Roll : Outcomes Multiplied

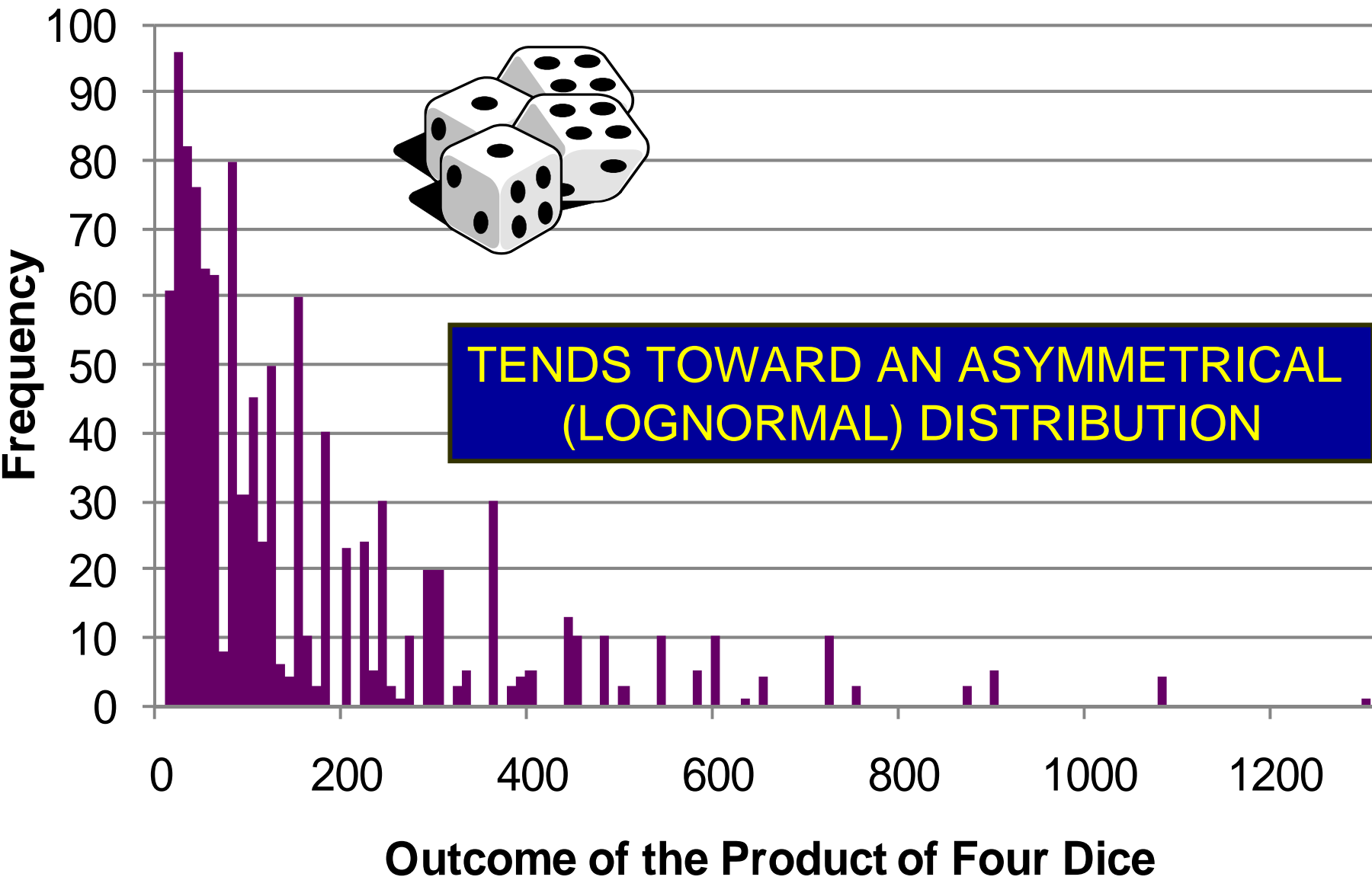


	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	8	10	12
3	3	6	9	12	15	18
4	4	8	12	16	20	24
5	5	10	15	20	25	30
6	6	12	18	24	30	36

# Product of Independent Distributions

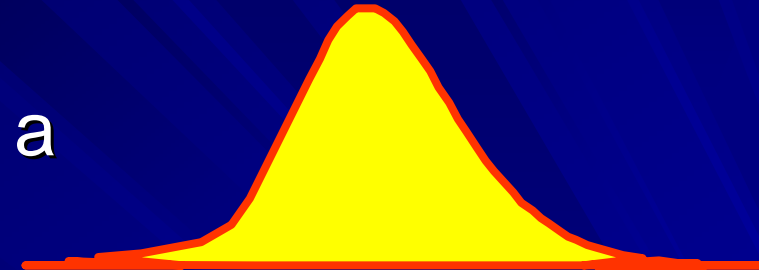


**TENDS TOWARD AN ASYMMETRICAL (LOGNORMAL) DISTRIBUTION**



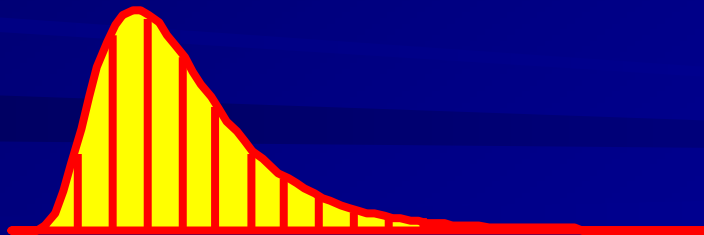
# Central Limit Theorem

**SUM** of a group of independent random variables tends towards a **NORMAL DISTRIBUTION**

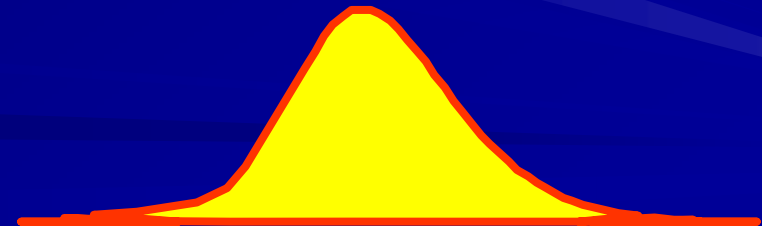


Linear Scale

**PRODUCT** of a group of independent random variables tends towards a **LOGNORMAL DISTRIBUTION**

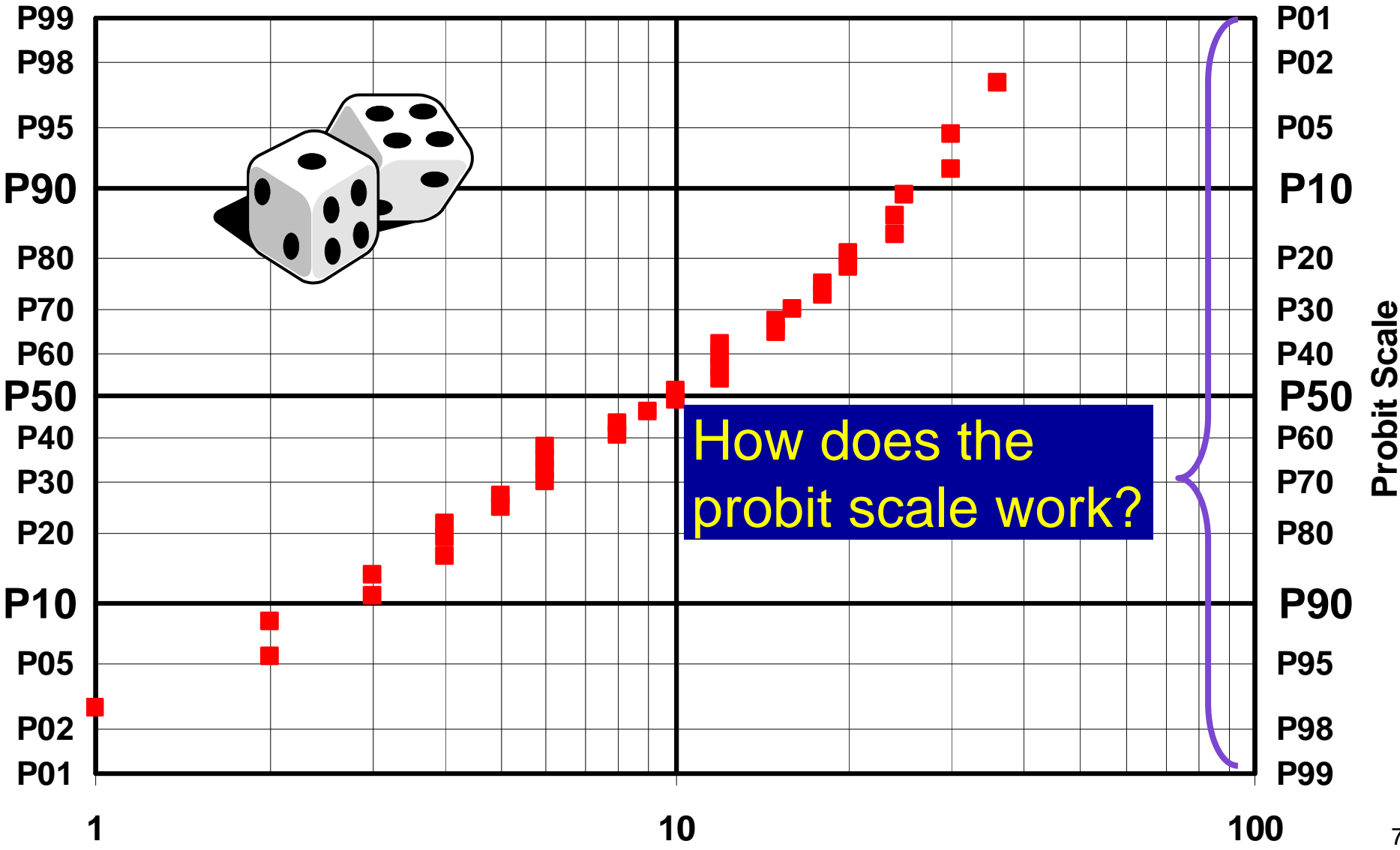
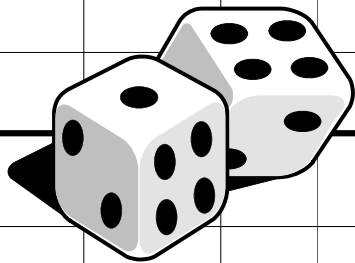


Linear Scale



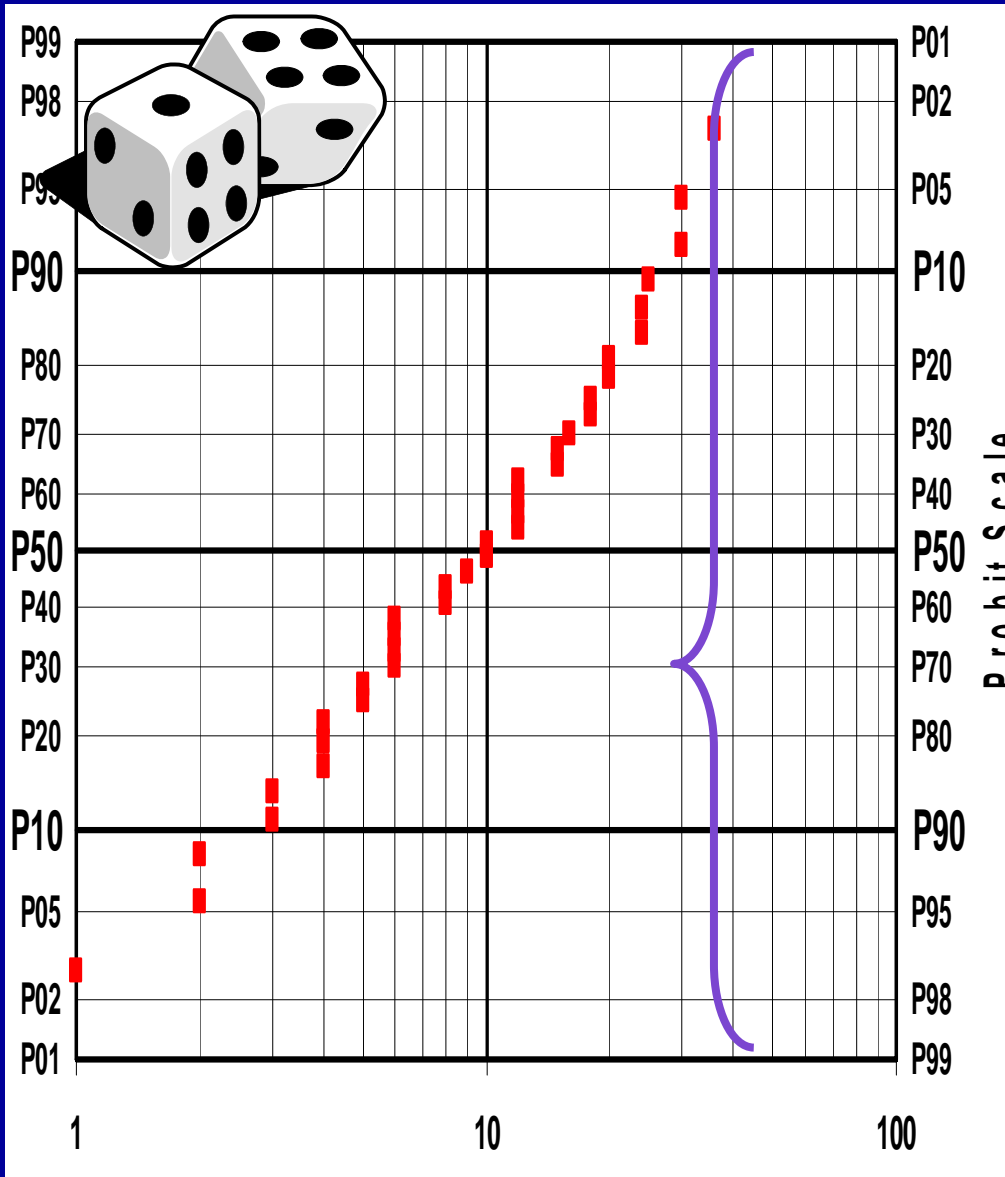
Log Scale

# Cumulative Probability Plot Product of Faces of Two Dice

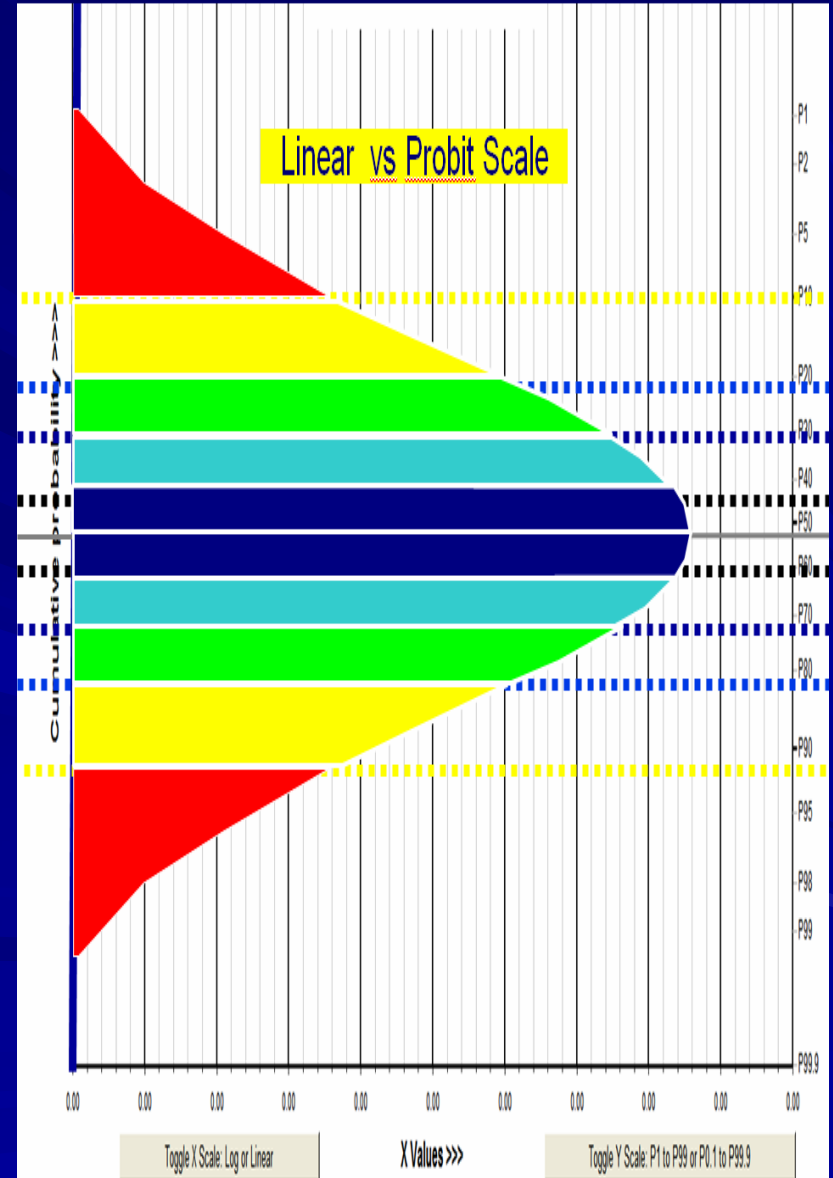


# How does the probit scale work?

Cumulative Probability Plot  
Product of Faces of Two Dice



Understanding the Cumulative  
Probability Plot  
Y Axis



# Plotting Conventions

- Definitions: %  $\geq$  ('GE') or %  $\leq$  ('LE')
- Evolving preference: %  $\geq$  ('GE')
  - Explorers think in terms of large discoveries
  - Consistent with SEC / SPE / WPC / AAPG guidelines
  - Commercial threshold truncations easier to apply
  - Less confusing for decision makers

In a Greater Than convention:

- P10 is the larger number
- P90 is the smaller number

# Exercise PB-4: Production Rate Distribution

Production Rate (Mcf/d): 850, 2500, 570, 1100,  
160, 333, 1333

Using the mid-point approach, build a field size distribution on log probability paper:

	<u>MCFD</u>	<u>Rank</u>	<u>%tile</u>
a. Arrange all fields by size, largest first and assign Rank	2500	1	7.2
b. Determine the Percentile Values	1333	2	21.4
	1100	3	35.7
	850	4	50
	570	5	64.3
c. Tabulate Field Sizes with %tiles	333	6	78.6
d. Plot the data pairs	160	7	92.9

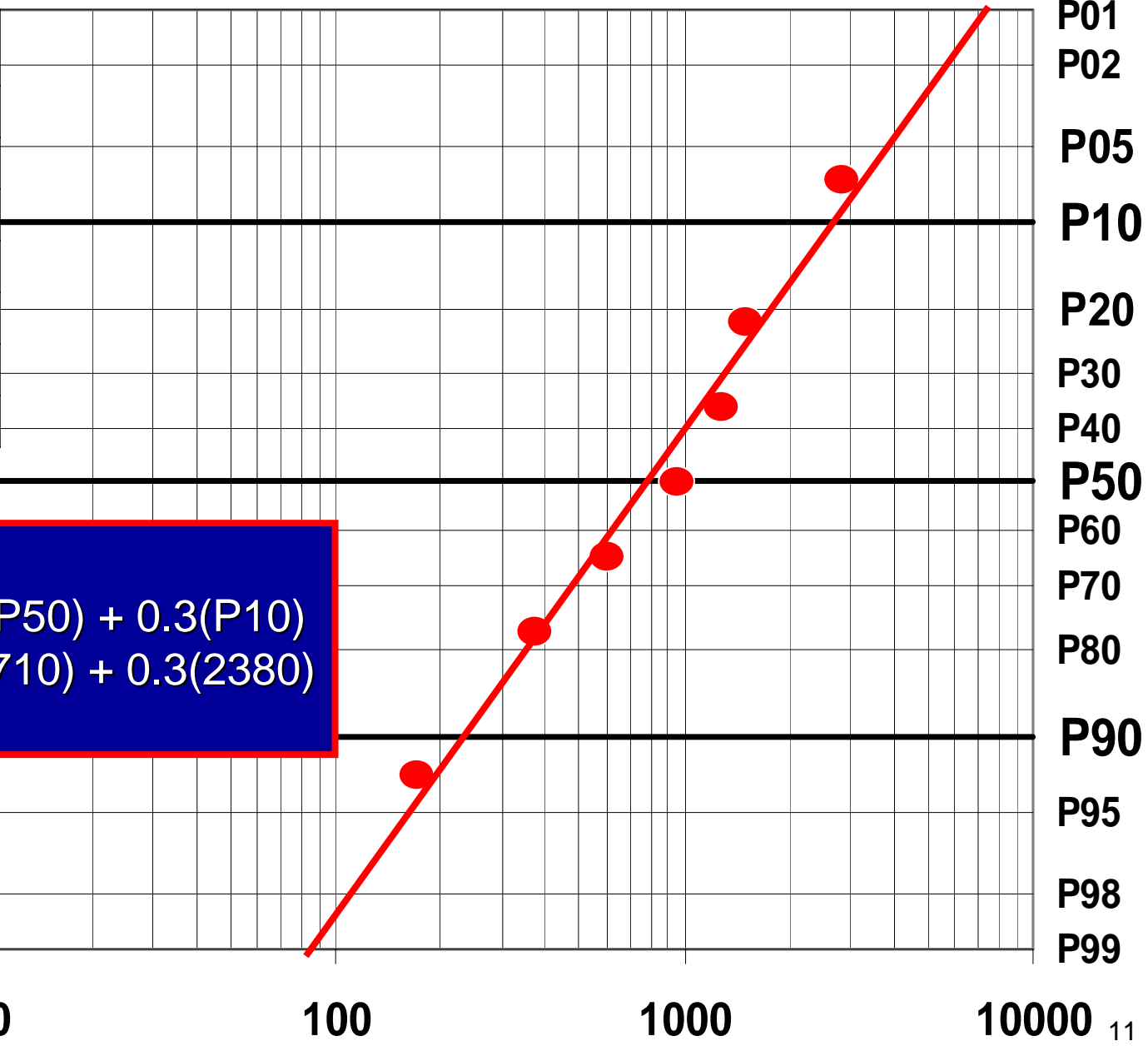
$$n = 7; \text{ Percentile} = 100/n * (\text{Rank} - 0.5)$$

# Rank Order the Rates & Determine the Percentile

MCF	Percentile
2500	7.2
1333	21.4
1100	35.7
850	50
570	64.3
333	78.6
160	92.9

Arith mean: 978

Swanson's Mean  
 =  $0.3(P90) + 0.4(P50) + 0.3(P10)$   
 =  $0.3(215) + 0.4(710) + 0.3(2380)$   
 = 1063





**What if you have drilled 100 wells in a resource play, where the mean EUR is 3.5 BCF and the range is between 1 and 6 BCF - - -**

Will you bet your house, that the next well (you only get one chance) drilled in the resource play will be equal to, or greater than the average 3.5 BCF per well?

*Not likely?*

*Why not?*

*What EUR would you be willing to Bet and be reasonably certain.*

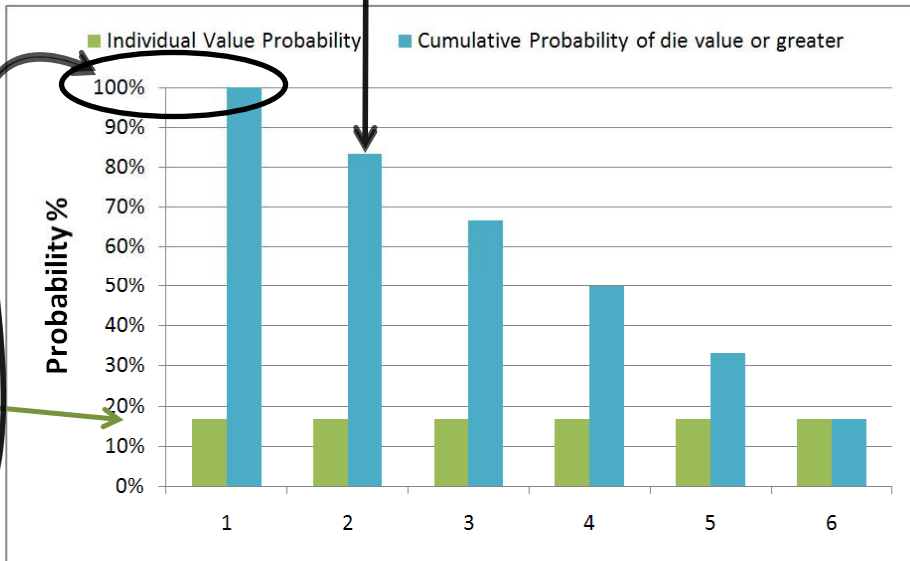
# Aggregation -- One Die Distribution



Chance of rolling a 1 or 2 or 3 etc. is 83%

Chance of rolling a 2 or Greater is 83%

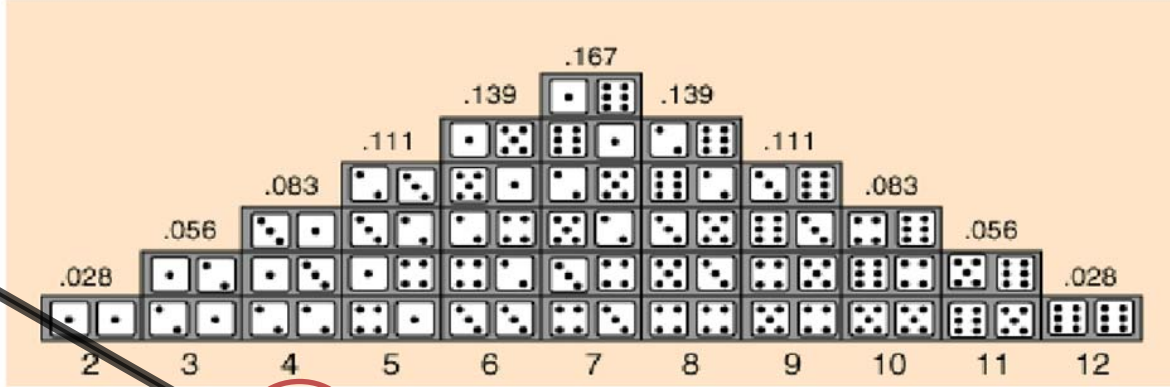
Die Values	1	2	3	4	5	6	AVG
Die Values							3.5
Frequency	1/6	1/6	1/6	1/6	1/6	1/6	
Probability %	17%	17%	17%	17%	17%	17%	
Cum Frequency	100%	83%	67%	50%	33%	17%	



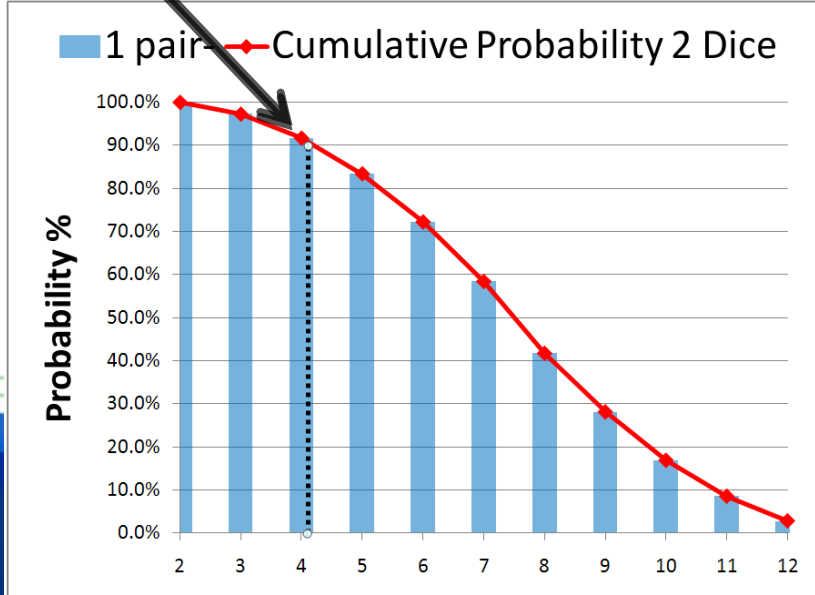
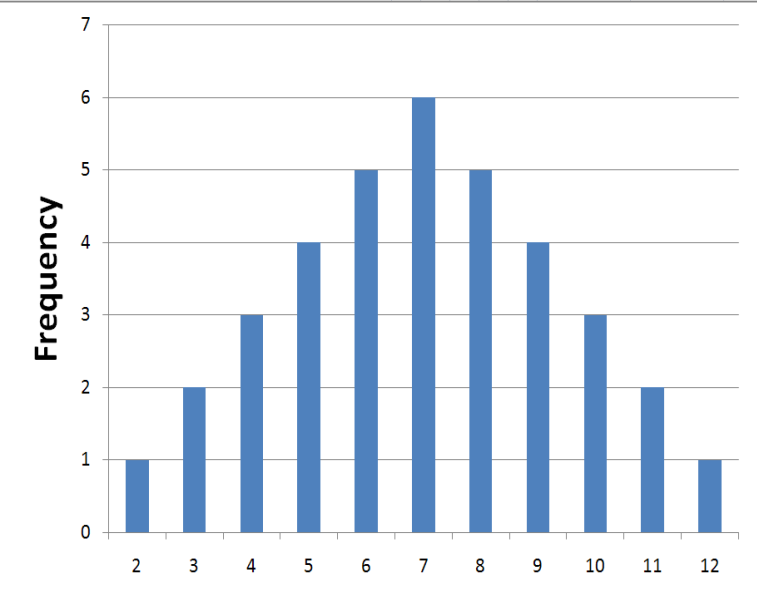


# Aggregation Example - Two Dice Distribution

Criteria = 2



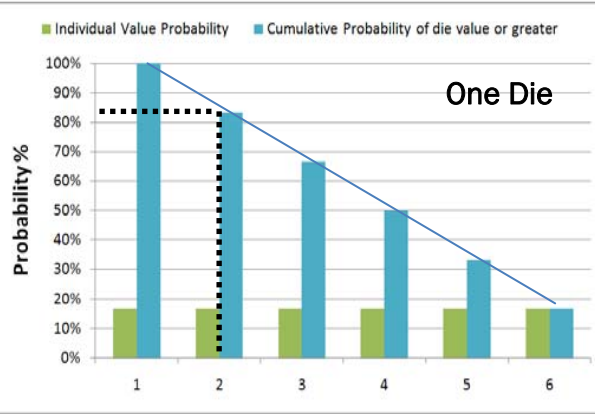
Possible Dice Values	2	3	4	5	6	7	8	9	10	11	12
Frequency	1	2	3	4	5	6	5	4	3	2	1
Cum Frequency	1	3	6	10	15	21	26	30	33	35	36
Cum Probability	100%	97.2%	91.7%	83.3%	72.2%	58.3%	41.7%	27.8%	16.7%	8.3%	2.8%



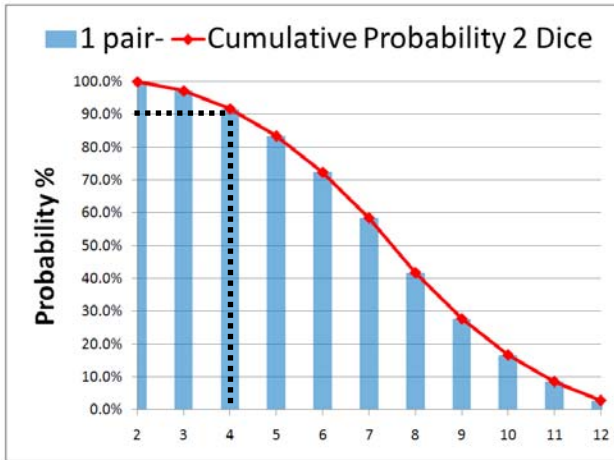
# Aggregation and Probability



If we hold the criterion value per Die at 2 or Greater, What happens to the Probability (confidence)

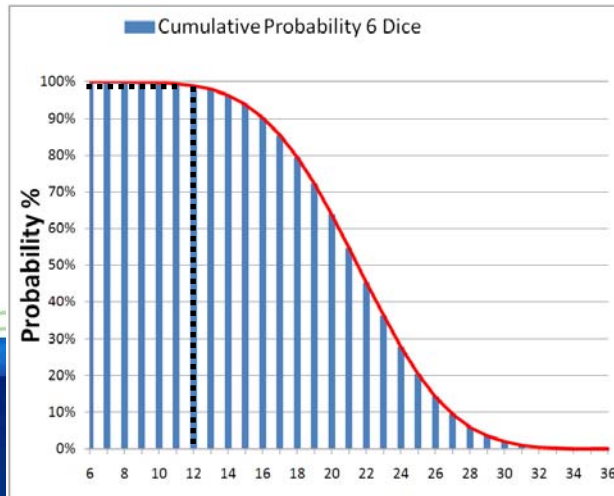


~P83% for value of 2 or Greater



~P90% for 2 Dice X 2 Dice = 4

**CONCLUSION:** confidence criterion must change with added distributions, otherwise "high confidence" becomes increasingly conservative

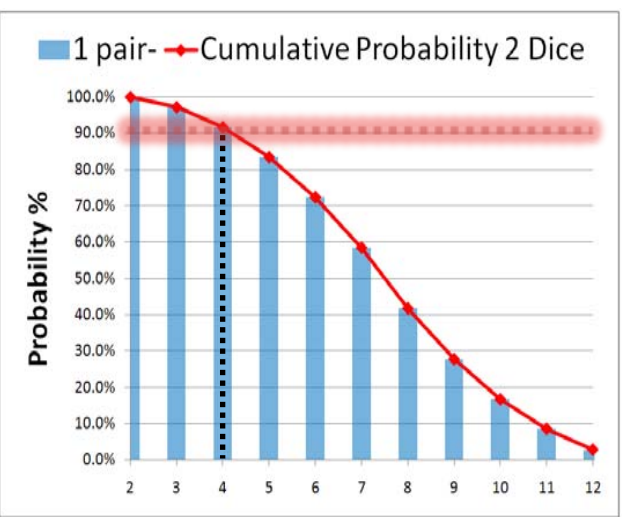


~P99.8% 2 X 6 Dice = 12

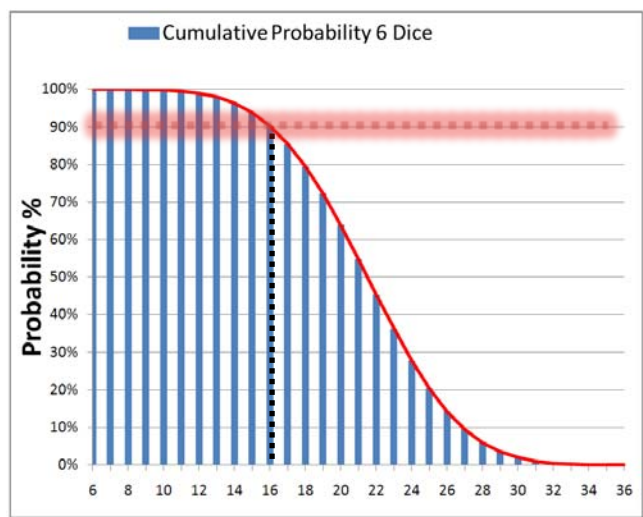


# Aggregation – What is the P90 Equivalent of One Die for various sample sizes

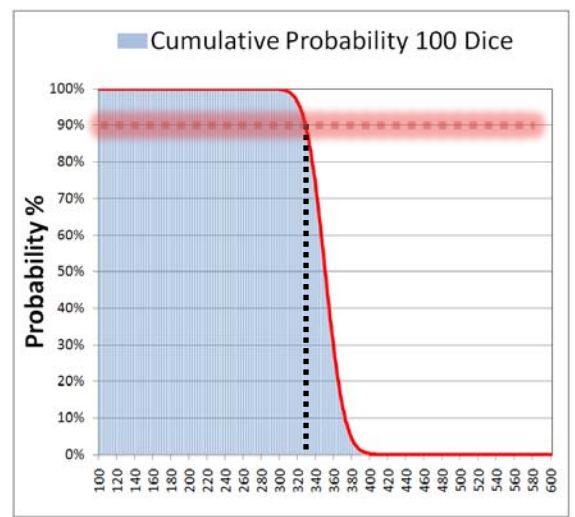
The correct question: What value represents the P90 or Proved of various Distributions?



4.2 & 2 Dice = 2.1 per Die



16.1 & 6 Dice = 2.7 per Die

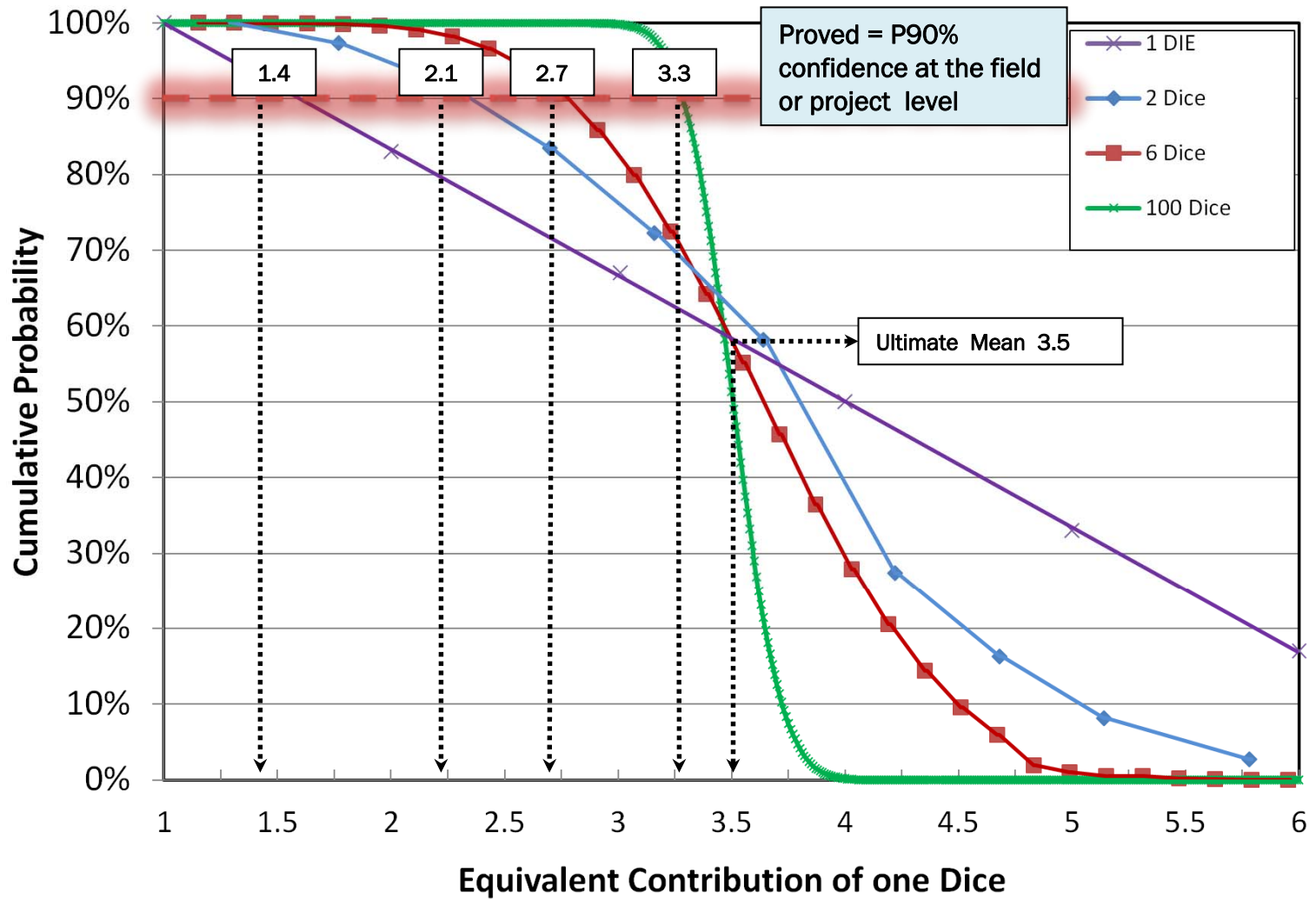


330 & 100 Dice = 3.3 per Die



# P90 comparison for 1, 2, 6, and 100

## Example Aggregation : 4 Distributions





# Aggregation: Proved?

## FIELD PDP Performance Information:

Existing Wells	100
AVG EUR	3.5
Range of EUR	1-6
P10/P90	4

What can be booked as Proved for each case?

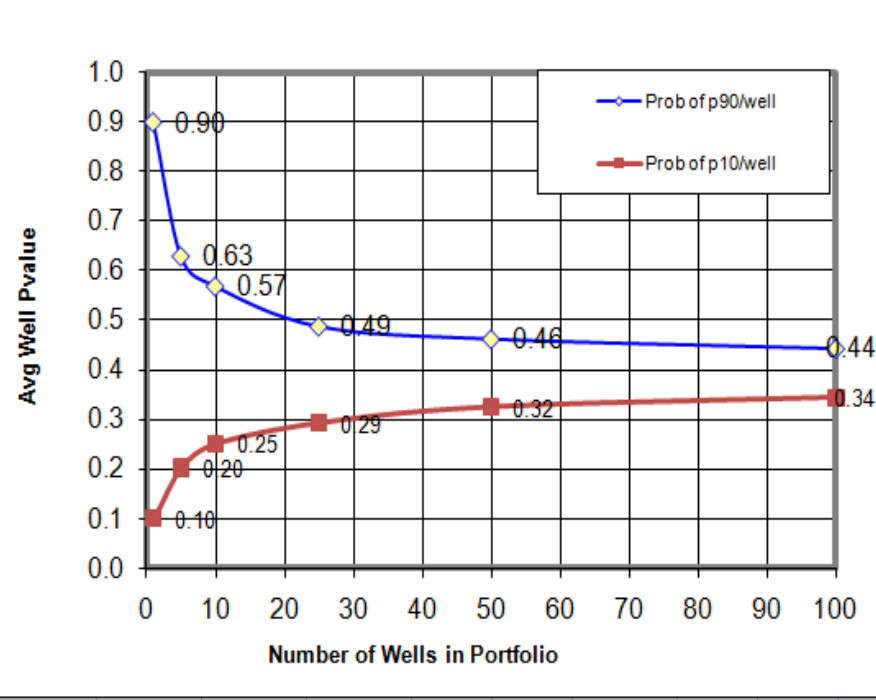
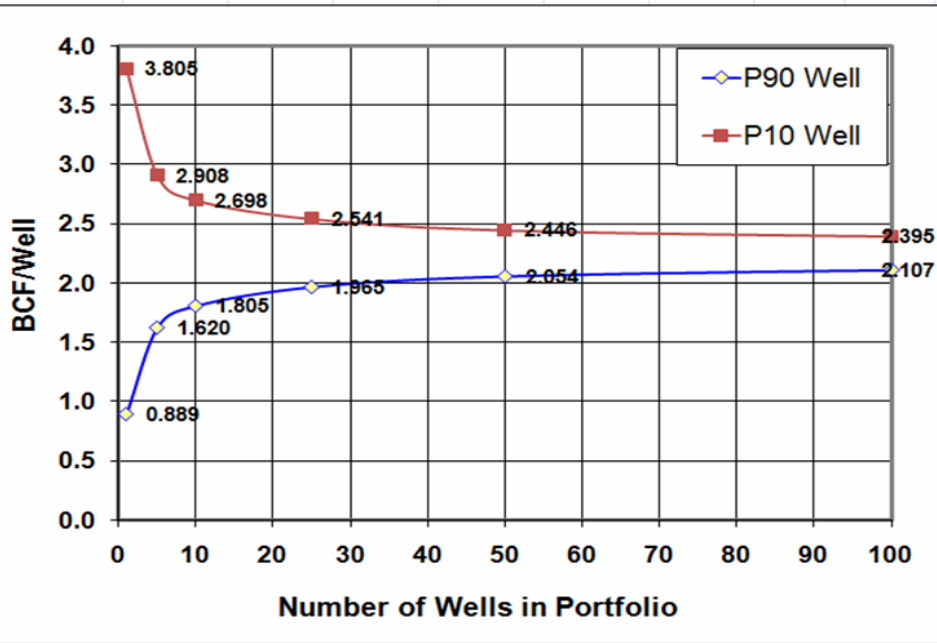
<u>PUD Locations</u>	1	2	6	100
<u>EUR per well</u>	1.4	2.1	2.7	3.3



# Aggregation Happens Quickly

P90 Bookable value for various portfolio well counts where the mean is 2.2

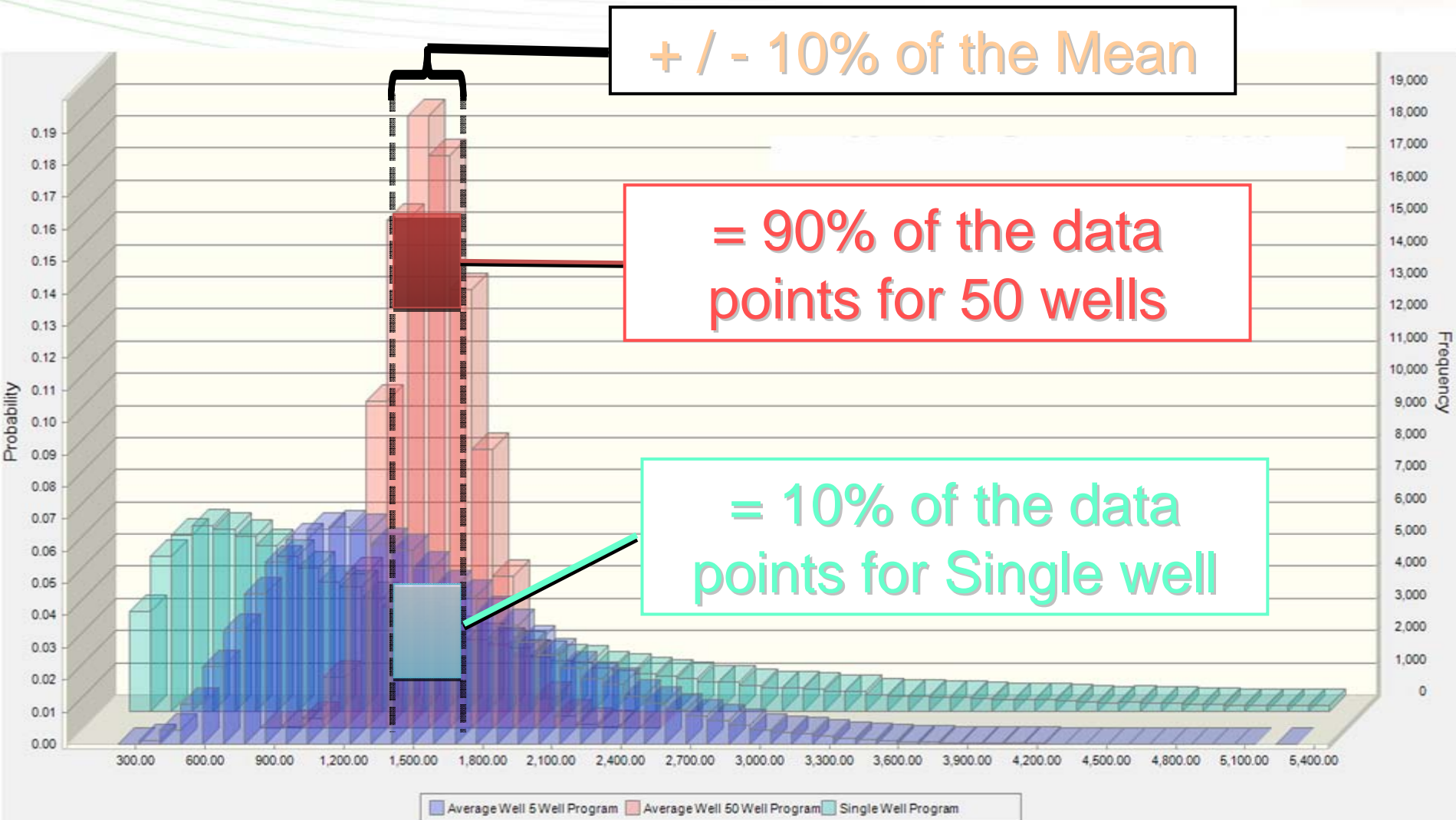
Probability that a P90 well would book at for various portfolio well counts



— Monte Carlo simulated P10 and P90 for different Portfolio well counts

# Aggregation Histograms

## Variability around the Mean

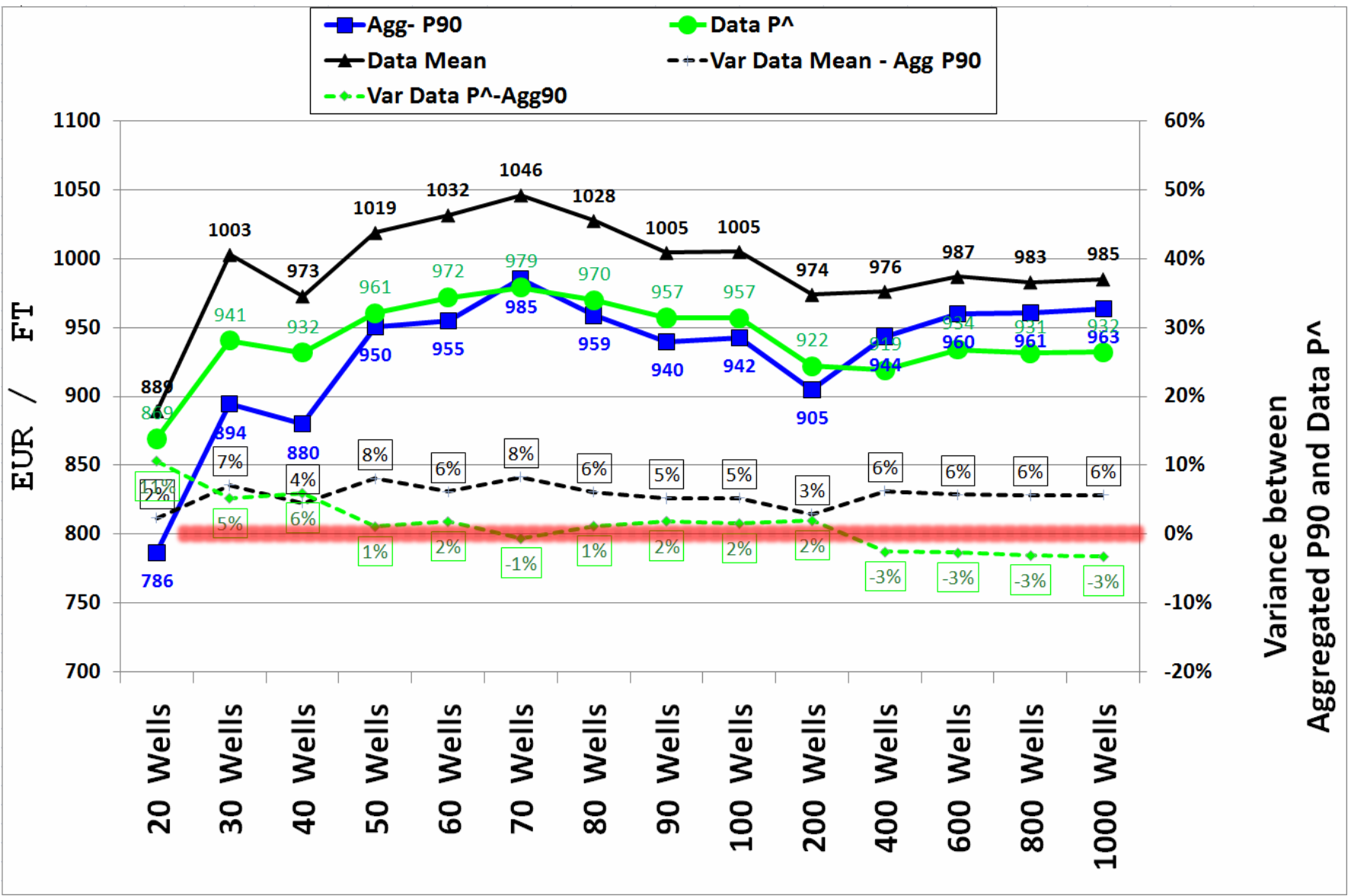


Aggregation impact of 1, 5 and 50 wells on a  $P_{10}/P_{90}$  ratio of 10 Type Curve

# The P^ Approximation Concept

- A simple first approximation of the Aggregated portfolio P90 of a large number of PUD locations, which will be represented by the same Type Curve, is the P^ of the distribution.
- Where:  **$P^ = (P50 + PMean)/2$**
- P^ is a good approximation method for well counts in excess of 30.
- P^ is slightly optimistic for well counts less than 70 to 100 for a P10/P90 Ratio of 4 to 8 respectively.
- P^ is slightly pessimistic for well counts in excess of 70 to 100 for a P10/P90 Ratio of 4 to 8 respectively.

# P^ as a Proxy For The Aggregated P90



# $P^{\wedge}$ as a Proxy For The Aggregated P90

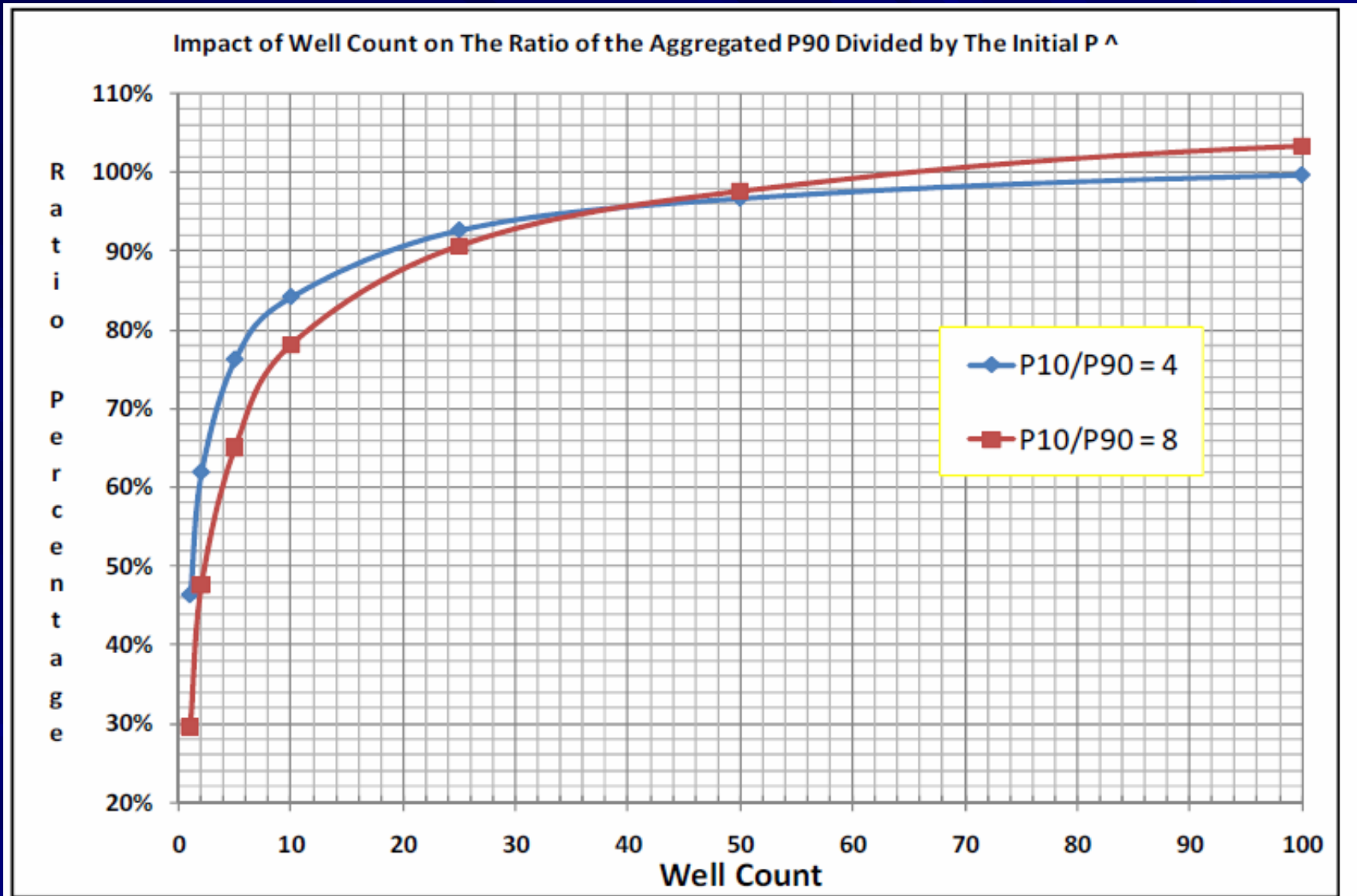
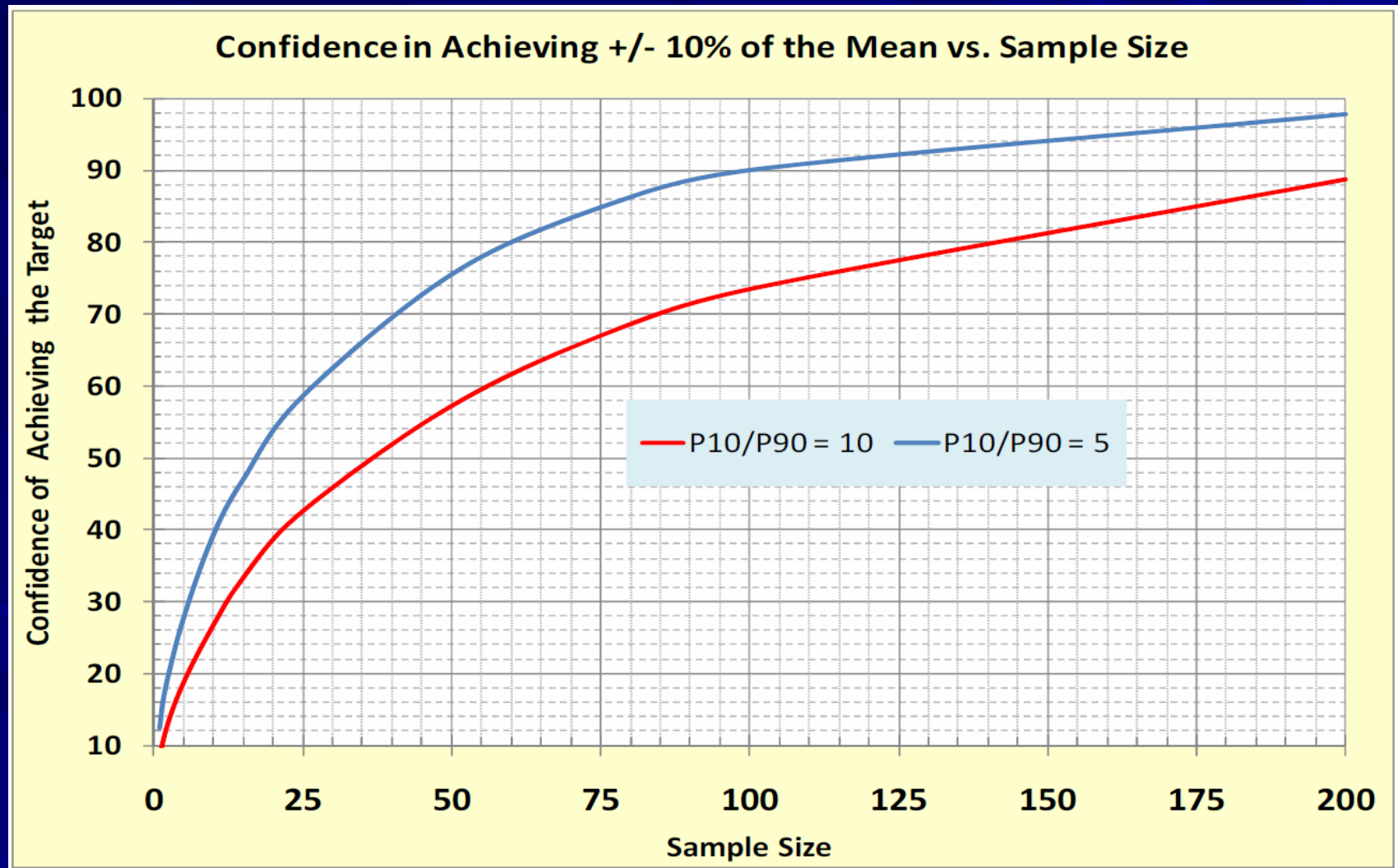


Fig. 2.31 – Ratio of Aggregated  $P_{90}/P^{\wedge}$  Vs. Well Count

# Confidence in Achieving +/- 10% of the Mean as a function of Sample Size



# What Distributions Are Horizontal Wells?

- Lognormal for each stage
- Tending towards Normal as we add stages
  - ❖ However each zone impacts the total results are partially dependent
- Deviations from Normal and Lognormal at both extremes.
  - ❖ High end due to extensive pre-existing fracture networks
  - ❖ High end due to no flow boundaries
  - ❖ Low end due to multiple stages of fracturing; hence an unlikely very poor outcome