Multivariate Modeling: Probability and Statistics

Lecture outline . . .

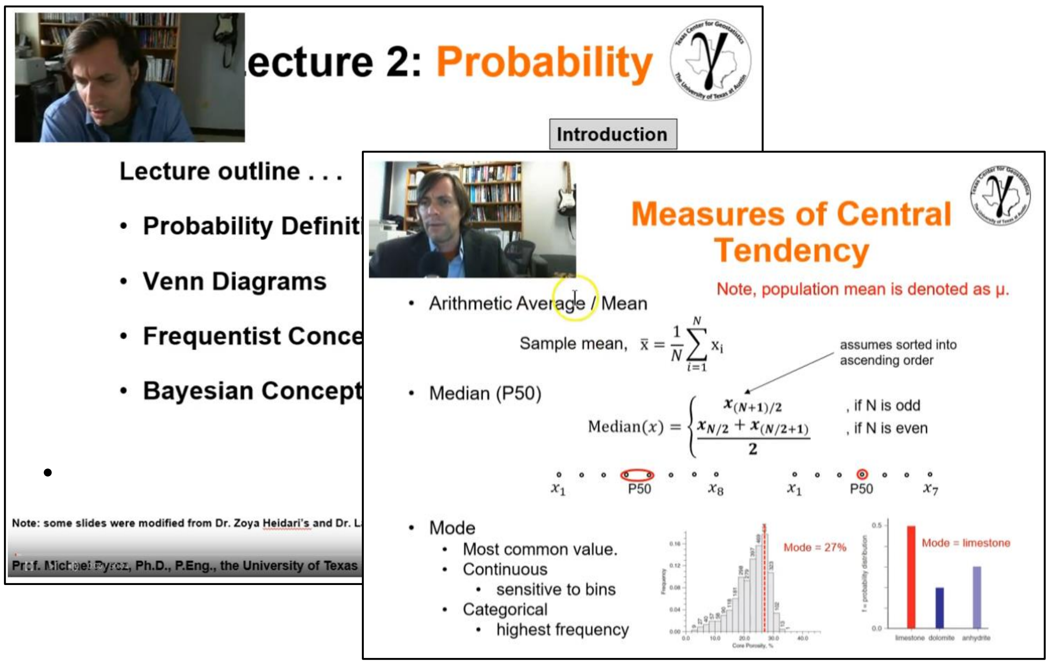
- Probability in Subsurface Modeling
- Frequentist Concepts
- Bayesian Concepts

Instructor: Michael Pyrcz, the University of Texas at Austin

Multivariate Modeling: Probability and Statistics

Other Resources:

- **Lectures recorded on YouTube.**



Probability and statistics lectures on YouTube.

- **Worked out examples on GitHub in Excel and Python**

Introduction

Fundamental Concepts

Probability

Data Prep / Analytics

Spatial Continuity / Prediction

Multivariate Modeling

Uncertainty Modeling

Machine Learning

Multivariate Modeling: Probability and Statistics

Lecture outline . . .

- **Probability in Subsurface Modeling**

Introduction

Fundamental Concepts

Probability

Data Prep / Analytics

Spatial Continuity / Prediction

Multivariate Modeling

Uncertainty Modeling

Machine Learning

Instructor: Michael Pyrcz, the University of Texas at Austin

Probability and Statistics
What should you learn from this lecture?

- **Fundamentals of Statistics and Probability**
  - **Fundamentals of Probability**
    - » **Basic Definitions and Rules**
    - » **Venn Diagram**
    - » **Conditional Probability**
    - » **Probability tree**
    - » **Bayes' Theorem**
    - » **Applications of Probability in Decision Making**

Probability Supports Decision Making

For example:

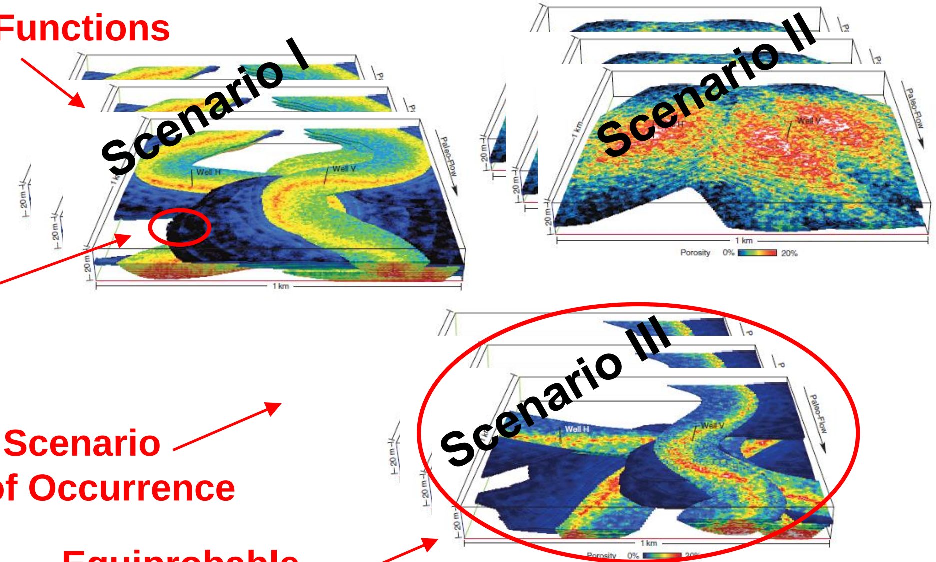
- What is the probability that a well is a success? – *drill the well*
- What is the probability that a valve has a crack? – *replace the valve*
- What is the probability that a seismic survey finds a reservoir? – *acquire the seismic*
- What is the probability that a reservoir seal will fail? – *inject the CO2*

Most of our decisions involve uncertainty:

- By quantifying probability we can make better decisions.

Probability in Modeling Workflows

Model Parameter
Probability Density Functions



Local Probability
of Occurrence

Discrete Scenario
Probability of Occurrence

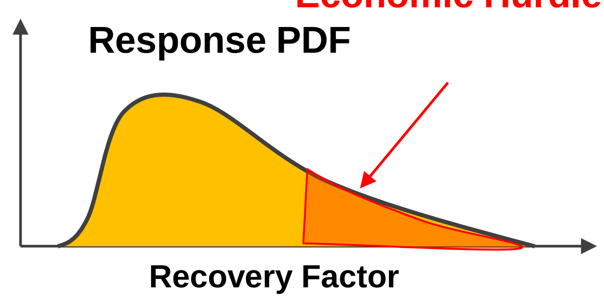
Equiprobable
Realizations

Geostatistical Subsurface Modeling

1. The entire workflow is based on probability (and statistics).
2. We must understand probability and statistics!
3. Let's make sure we are on the same page.

Transfer Function

Probability of Exceedin
Economic Hurdle



Probability Definitions
What is Probability?
Frequentist Approach

Measure of the likelihood that an event will occur. For random experiments and well-defined settings (such as coin tosses):

$$\text{Prob}(A) = \text{P}(A) = \lim_{n\to\infty}\left(\frac{n(A)}{n(\Omega)}\right)$$

frequentist approach to probability is the limit of relative frequency over a large number of trails.

where:

n(A) = number of times event A occurred
n(Ω) = number of trails

Example: Possibility of drilling a dry hole for the next well, encountering sandstone at a location ($\mathbf{u}_\alpha$), exceeding a rock porosity of 15% at a location ($\mathbf{u}_\alpha$).

Probability Definitions
What is Probability?
Bayesian Approach

Measure of the likelihood that an event will occur. For any occurrence:

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

Bayesian approach probability is interpreted as reasonable likelihood representing a state of knowledge or as quantification of a personal belief
Specify a prior and update with new information.

where:

$P(A)$ = prior
$P(B|A)$ = likelihood

$P(B)$ = evidence
$P(A|B)$ = posteriori

Example: Given a prior probability of 40%, a likelihood of 10% and evidence term of 20% the updated posteriori is 20%.

Multivariate Modeling: Probability and Statistics

Lecture outline . . .

- **Frequentist Concepts**

Introduction

Fundamental Concepts

Probability

Data Prep / Analytics

Spatial Continuity / Prediction

Multivariate Modeling

Uncertainty Modeling

Machine Learning

Instructor: Michael Pyrcz, the University of Texas at Austin

Probability Definitions
What is Probability?

We will start with Frequentist notions and then move to Bayesian approaches.

Knowledge of both is essential as there are many classes of problems that can only be addressed practically with Frequentist or Bayesian approaches.

- We need both frequentist and Bayesian frameworks

We build up to Bayesian Updating with frequentist concepts but we accept the role of belief and updating with new evidence.

Probability Concepts
Venn Diagrams

Venn Diagrams are a tool to communicate probability

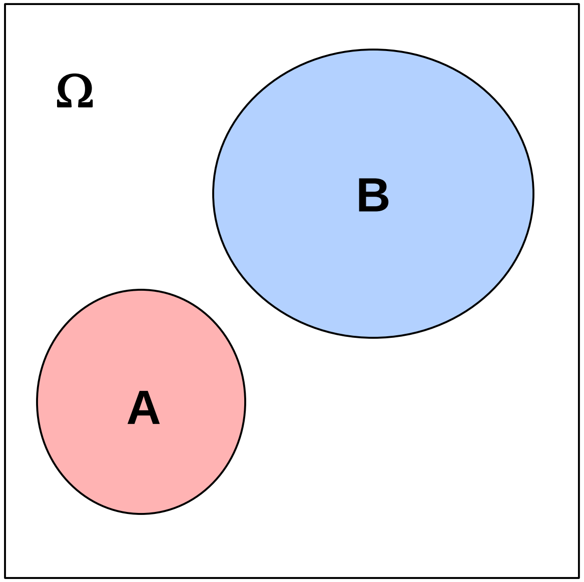
Experiments (Sampling) (J): Establishment of conditions that produce an outcome.

Simple Event (x): A single outcome of an experiment.

Event (A, B, …): Collection of simple events.

Occurrence of A: A has occurred if the outcome of experiment (sampling) belongs to it.

Sample Space (Ω): Collection of all possible events.



Venn Diagram – illustration of events and relations to each other.

What do we learn from a Venn diagram?

- size of regions = probability of occurrence
- overlap = probability of joint occurrence
- excellent tool to visualize marginal, joint and conditional probability.

Probability Definitions
Venn Diagram Example

Experiments (Sampling) (J):

- Facies determined from a set of well cores (N=3,000 measures at 1 foot increments)

Sample Space (Ω):

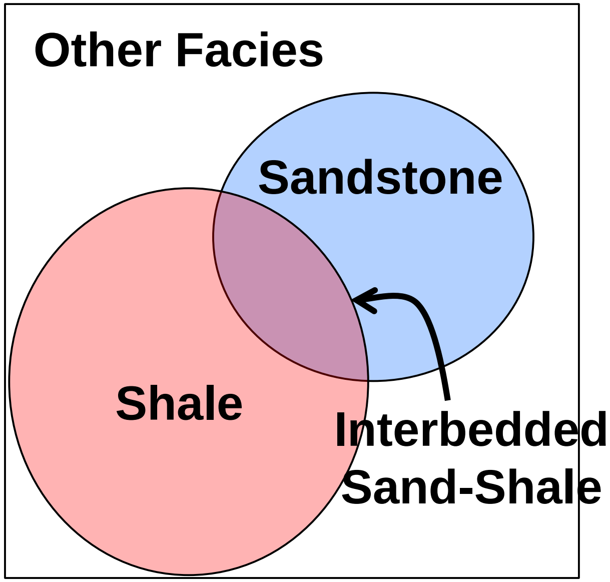
- Facies for the N=3,000 core measures

Event (A, B, …):

- Facies = {Sandstone, Interbedded Sandstone and Shale, Shale and other facies}

Venn Diagram Tells Us About Probability:

- Prob{Other Facies} > Prob{Shale} > Prob{Sandstone} > Prob{Interbedded} = Prob{Shale and Sandstone}
- Prob{Sandstone and Shale given Sandstone } < Prob{Sandstone}



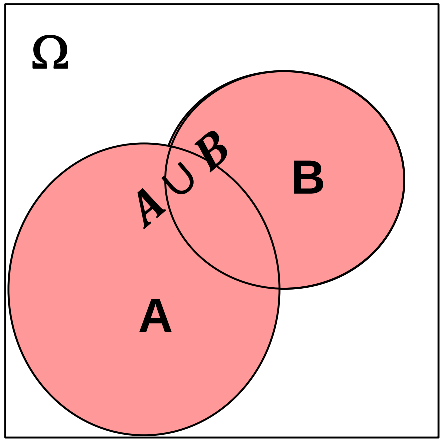
Venn Diagram – illustration of events and relations to each other.

Probability Definitions
Probability Operators

Union of Events:

- All outcomes in the sample space that belong to either event A or B

$$A \cup B = \{x: x \in A \; or \; x \in B\}$$

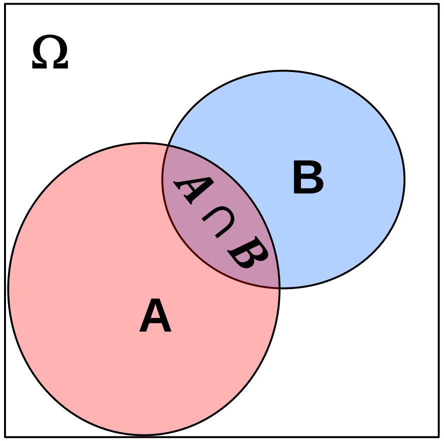


Venn Diagram – illustrating union.

Intersection of Events:

- All outcomes in the sample space that belong to both events A and B

$$A \cap B = \{x: x \in A \; and \; x \in B\}$$

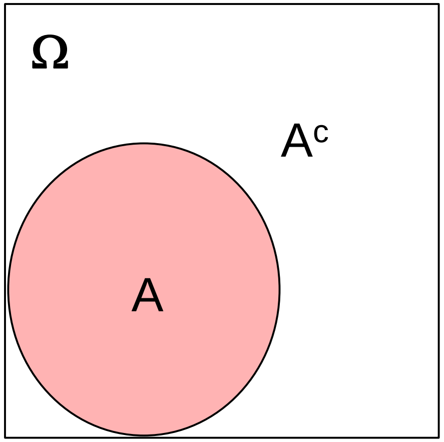


Venn Diagram – illustrating intersection.

Probability Definitions
Probability Operators

Complementary Events: A^c

- All outcomes in the sample space that do not belong to A

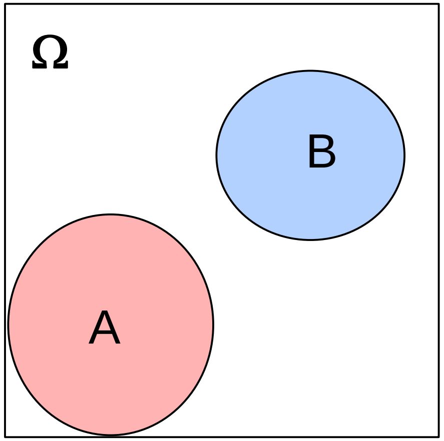


Venn Diagram – illustrating complementary events.

Mutually Exclusive Events:

- The events that do not intersect or do not have any common outcomes

$$A \cap B = \emptyset \rightarrow \text{Null Set}$$



Venn Diagram – illustrating mutually exclusive.

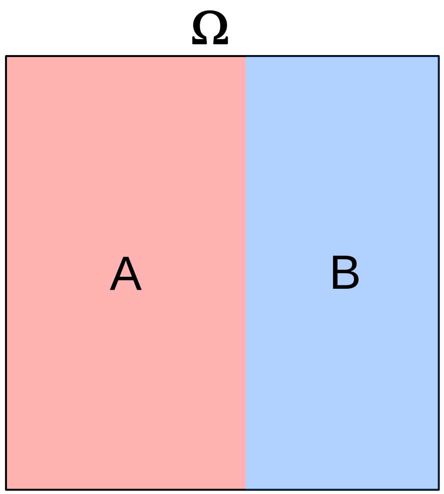
Probability Definitions
Probability Operators

Exhaustive, Mutually Exclusive Sequence of Events:

- The sequence of events whose union is equal to the sample space

$$A_1 \cup A_2 \cup \ \ldots \cup A_n = \Omega$$

- For example, all the samples are either A or B



Venn Diagram – illustrating exhaustive events.

Probability Definitions Now We Refine Probability

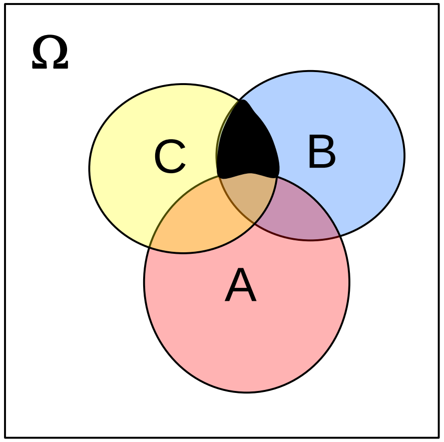
where:

$$\mathrm{Prob}(A) = \mathrm{P}(A) = \lim_{n\to\infty}\left(\frac{Area(A)}{Area(\Omega)}\right)$$

Area(A) = area of A / total area = P(A)
Area(Ω) = total area / total area = probability of any possible outcome = P(Ω) = 1.0

Example: Possibility of drilling a dry hole for the next well (A^C), encountering sandstone at a location ($\mathbf{u}_\alpha$)(B), exceeding a rock porosity of 15% at a location ($\mathbf{u}_\alpha$)(C).

Prob($\boldsymbol{A^C} \cap \boldsymbol{B} \cap \boldsymbol{C}$) = Area($\boldsymbol{A^C} \cap \boldsymbol{B} \cap \boldsymbol{C}$) / Area(Ω)



Probability Definitions
Probability Concepts

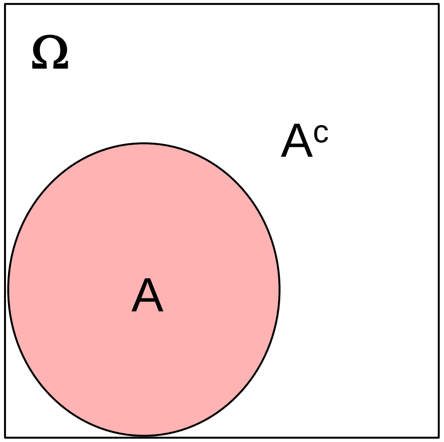
Non-negativity, Normalization:

- Fundamental probability constraints

  - Bounded $0 \leq P(A) \leq 1$
  - Closure $P(\Omega) = 1$
  - Null Sets $P(\phi) = 0$

Complimentary Events:

  - Closure $P(A^c) + P(A) = 1$



Venn Diagram – illustrating complementary events.

Probability Definitions
Probability Concepts

The Addition Rule:

$$P(A \cup B) = \mathrm{P(A) + P(B) - P}(A \cap B)$$

Must account for the intersection!

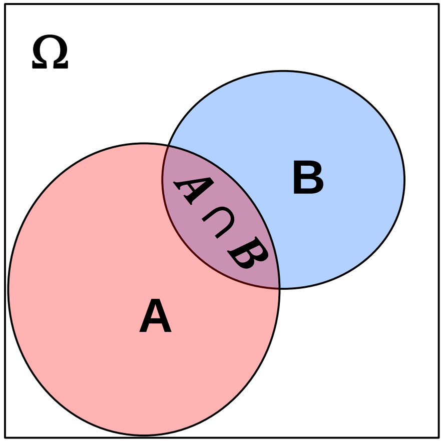
If mutually exclusive events:

$$A_i \cap A_j = \emptyset, \forall\, i \neq j$$

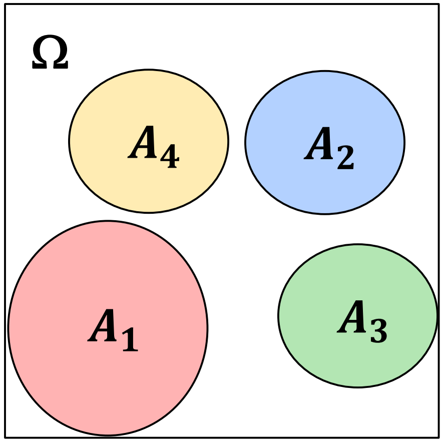
then,

$$\mathrm{P}\left(\bigcup_{i=1}^{k} \mathrm{A}_i\right) = \sum_{i=1}^{k} \mathrm{P}(\mathrm{A}_i)$$

no intersections to account for.



Venn Diagram – illustrating intersection.



Venn Diagram – illustrating intersection.

Probability Definitions Hands-on Addition Rule Example

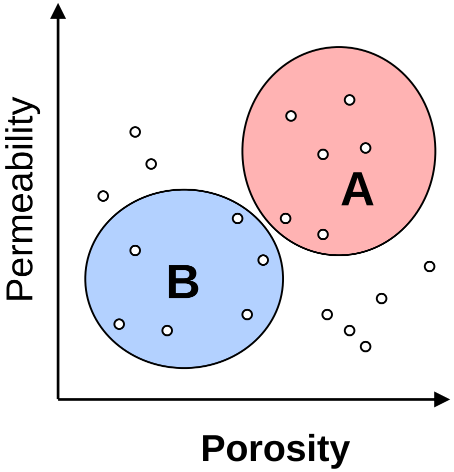
Calculate the following probabilities for event A and B: Note Event A: Sandstone and Event B: Shale

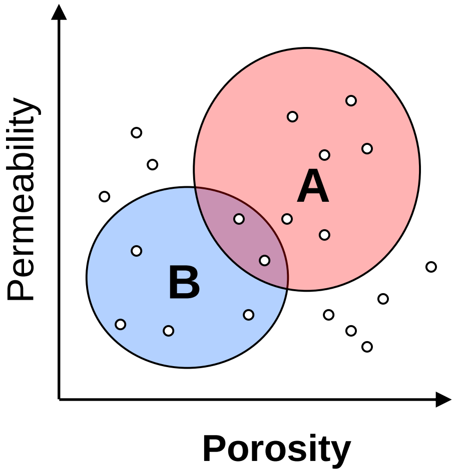
$$P(A) =$$

$$P(B) =$$

$$P(A \cap B) =$$

$$P(A \cup B) =$$





$$P(A) =$$

$$P(B) =$$

$$P(A \cap B) =$$

$$P(A \cup B) =$$

Recall: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ Hint: count points, don't calculate area.

Probability Definitions Addition Rule Example

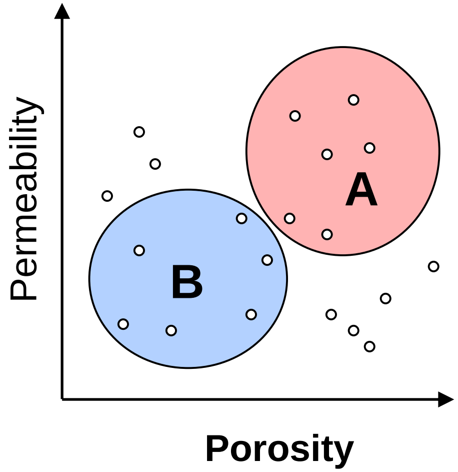
Calculate the following probabilities for event A and B: Note Event A: Sandstone and Event B: Shale

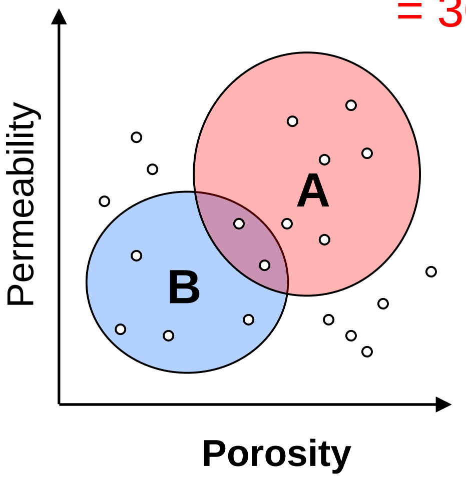
$$P(A) = {}^{6}/_{20} = 30\%$$

$$P(B) = {}^{6}/_{20} = 30\%$$

$$P(A \cap B) = {}^{0}/_{20} = 0\%$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 30\% + 30\% - 0\% = 60\%$$





$$P(A) = {}^{8}/_{20} = 40\%$$

$$P(B) = {}^{6}/_{20} = 30\%$$

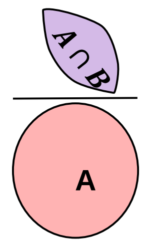
$$P(A \cap B) = {}^{2}/_{20} = 10\%$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 40\% + 30\% - 10\% = 60\%$$

Probability Definitions
Conditional Probability

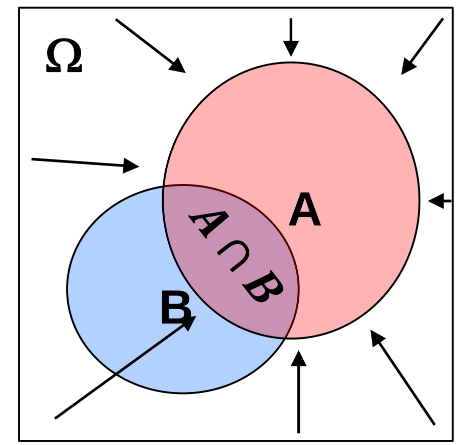
Probability of B given A occurred? $P(B \mid A)$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$



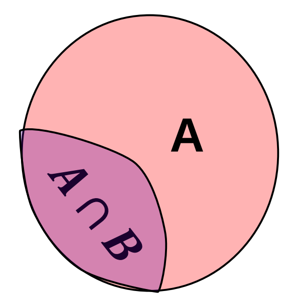
$P(A \cap B)$

$P(A)$



Conceptually we shrink space of possible outcomes.

A occurred so we shrink our space to only event A.



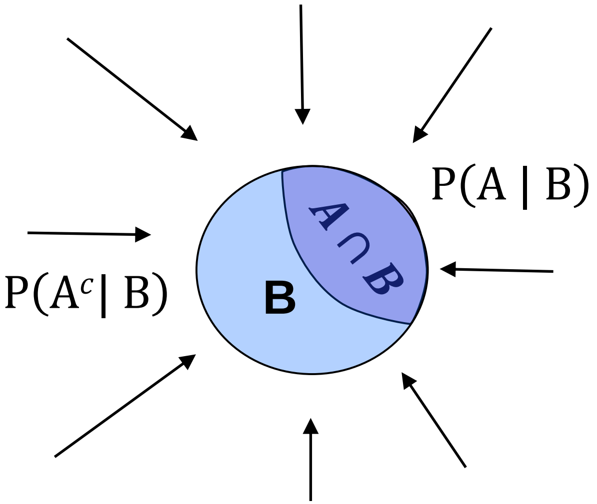
Probability Definitions
More on Conditional Probability

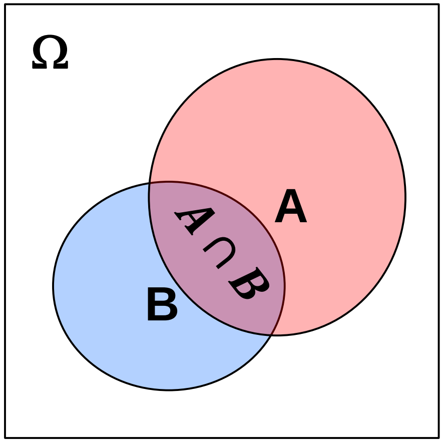
Other Relations with Conditional Probability

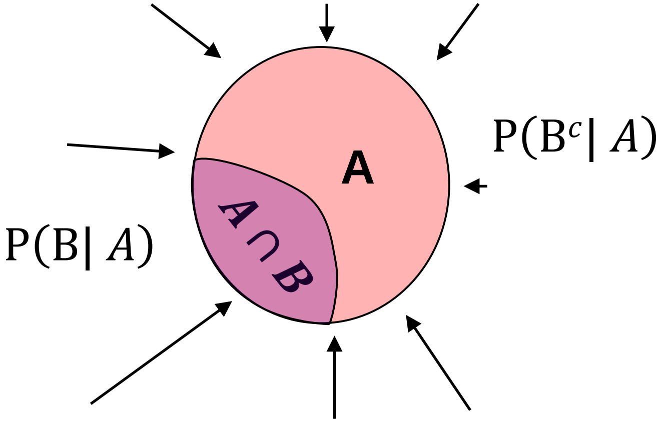
- Closure with conditional probabilities:

$$P(A \mid B) + P(A^c \mid B) = 1$$

$$P(B \mid A) + P(B^c \mid A) = 1$$







Probability Definitions
Conditional Probability Hands-on

Recall:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

For Case 1 calculate:

$P(A \mid B) =$

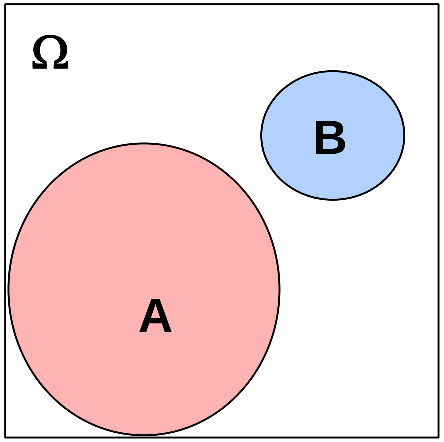
$P(B \mid A) =$

For Case 2 calculate:

$P(A \mid B) =$

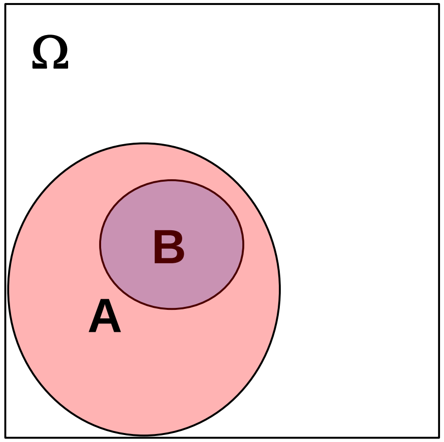
$P(B \mid A) =$

Case 1:



Venn Diagram – case 1.

Case 2:



Venn Diagram – case 2.

Probability Definitions
Conditional Probability Example

Recall:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

For Case 1 calculate:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{\emptyset}{P(B)} = \emptyset$$

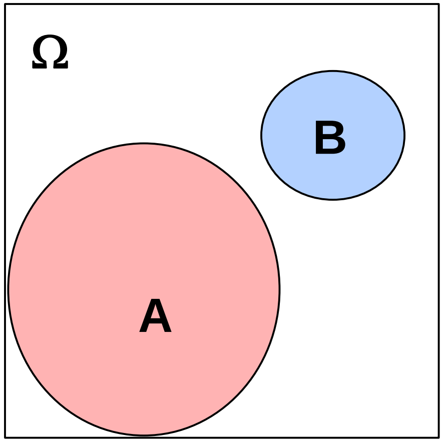
$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{\emptyset}{P(A)} = \emptyset$$

For Case 2 calculate:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A \cap B)}{P(B)} = 1, \text{since } P(A \cap B) = P(B)$$

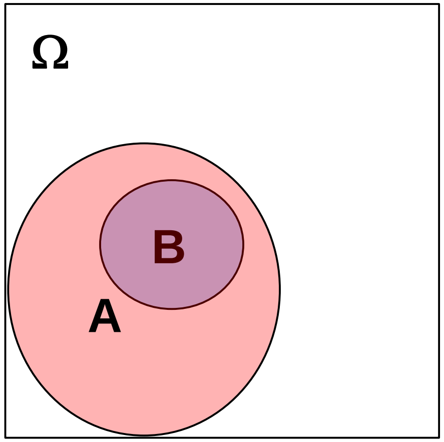
$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{P(B)}{P(A)}, \text{since } P(A \cap B) = P(B)$$

Case 1:



Venn Diagram – case 1.

Case 2:



Venn Diagram – case 2.

Probability Definitions
Conditional Probability Example

Question: Calculate the following probabilities for events A and B:

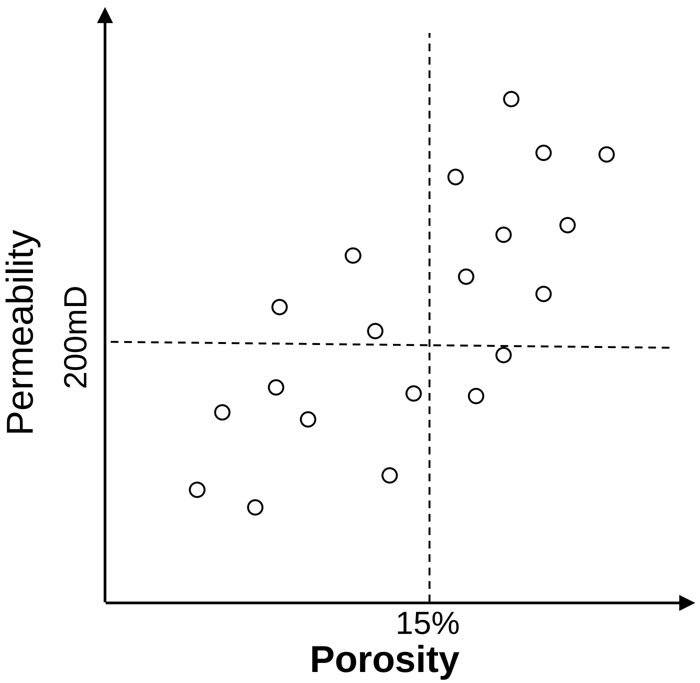
Event A: Porosity $> 15\%$
Event B: Permeability > 200 mD

For Case 1 calculate:

$P(A \mid B) =$

$P(B \mid A) =$

Bonus Question: How much information does event B tell you about event A and visa versa?



Recall $P(B \mid A) = \dfrac{P(A \cap B)}{P(A)}$

Probability Definitions
Conditional Probability Example

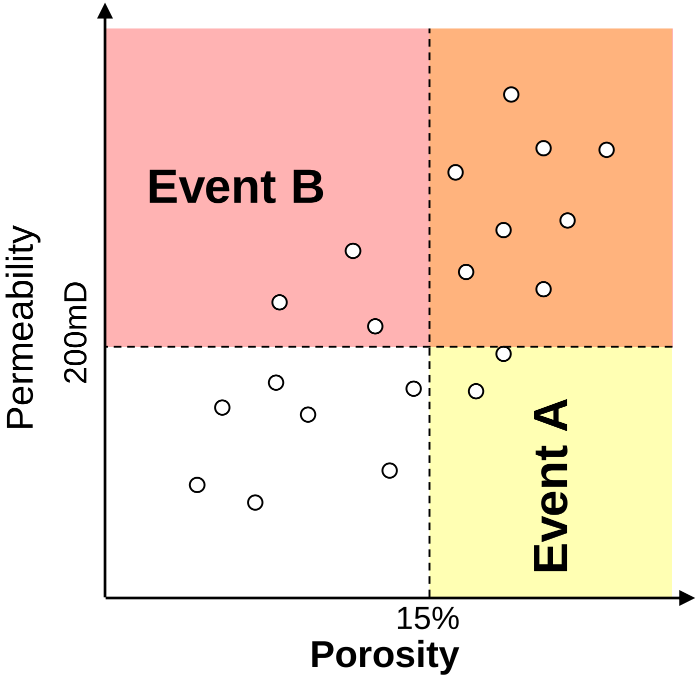
Question: Calculate the following probabilities for events A and B:

Event A: Porosity > 15%
Event B: Permeability > 200 mD

For Case 1 calculate:

$$P(A\,|\,B) = \frac{P(A \cap B)}{P(B)} = \frac{8/20}{11/20} = 8/11$$

$$P(B\,|\,A) = \frac{P(A \cap B)}{P(A)} = \frac{8/20}{10/20} = 8/10$$



Bonus Question: How much information does B tell you about A and visa versa?
P(A) = 10/20, P(A|B) = 8/11 Probability from 50% → 73%
P(B) = 11/20, P(B|A) = 8/10 Probability from 55% → 80%
We cannot work with A and B independently, they provide information about each other.

Probability Definitions

Conditional, Marginal and Joint Example

Joint Distribution:

$$f_{XY}(x,y)$$

Marginal Distribution:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y)dy$$

Conditional Distribution:

$$f_{X|Y}(x \mid y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

Table of Frequencies

Porosity (%) \ Fraction Shale (%)	10%	30%	50%	70%	90%
25%	1	1	0	0	0
20%	2	3	2	0	0
15%	1	2	2	1	0
10%	0	0	2	3	2
5%	0	0	1	1	1

Probability Definitions
Conditional Probability Example

Joint Distribution:

$$f_{XY}(x,y)$$

Marginal Distribution:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y)dy$$

Conditional Distribution:

$$f_{X|Y}(x \mid y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

Table of Joint Probabilities

Porosity (%) \ Fraction Shale (%)	10%	30%	50%	70%	90%
25%	4%	4%	0	0	0
20%	8%	12%	8%	0	0
15%	4%	8%	8%	4%	0
10%	0	0	8%	12%	8%
5%	0	0	4%	4%	4%

Probability Definitions
Conditional Probability Hands-on

Given these joint probabilities calculate the:

Marginal Distributions:

Vsh	10%	30%	50%	70%	90%
$f_{Vsh}(v_{sh}) =$					

Porosity	5%	10%	15%	20%	25%
$f_{\varphi}(\varphi) =$					

Conditional Distribution:

Vsh	10%	30%	50%	70%	90%
$f_{Vsh\|\varphi}(v_{sh}\| \varphi = 15\%) =$					

Table of Joint Probabilities

Porosity (%) \ Fraction Shale (%)	10%	30%	50%	70%	90%
25%	4%	4%	0	0	0
20%	8%	12%	8%	0	0
15%	4%	8%	8%	4%	0
10%	0	0	8%	12%	8%
5%	0	0	4%	4%	4%

Probability Definitions
Conditional Probability Hands-on

Given these joint probabilities calculate the:

Marginal Distributions:

$f_{Vsh}(v_{sh})$

Vsh	10%	30%	50%	70%	90%
$f_{Vsh}(v_{sh})$	16%	24%	28%	20%	12%

Porosity	5%	10%	15%	20%	25%
$f_{\varphi}(\varphi) =$	12%	28%	24%	28%	8%

Conditional Distribution:

Vsh	10%	30%	50%	70%	90%
	1/6	1/3	1/3	1/6	0

$$f_{Vsh|\varphi}(v_{sh}|\, \varphi = 15\%) = f_{Vsh,\varphi}(v_{sh}, \varphi = 15\%)/f_{\varphi}(\varphi = 15\%)$$

Table of Joint Probabilities

Porosity (%) \ Fraction Shale (%)	10%	30%	50%	70%	90%
25%	4%	4%	0	0	0
20%	8%	12%	8%	0	0
15%	4%	8%	8%	4%	0
10%	0	0	8%	12%	8%
5%	0	0	4%	4%	4%

Multivariate Modeling: Probability and Statistics

Lecture outline . . .

- **Bayesian Concepts**

Instructor: Michael Pyrcz, the University of Texas at Austin

Probability Definitions
Bayesian Statistics

Product Rule:

$$P(B \cap A) = P(A|B)\,P(B)$$

$$P(A \cap B) = P(B|A)\,P(A)$$

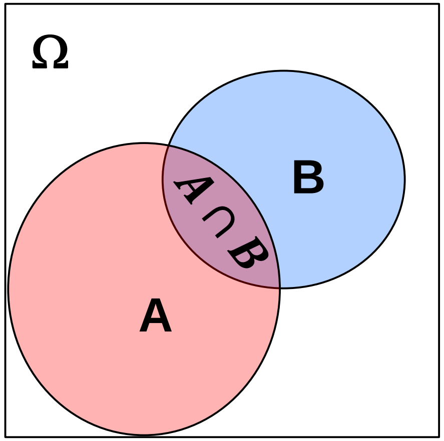
It follows that:

$$P(B \cap A) = P(A \cap B)$$

Therefore we combine two product rules, substitute:

$$P(A|B)\,P(B) = P(B|A)\,P(A)$$

We get Bayes' Theorem!



Venn Diagram – illustrating intersection.

Probability Definitions Bayesian Statistics

Bayesian Statistical Approaches:

- probabilities based on a degree of belief (expert experience) in an event
- updated as new information is available
- solve probability problems that we cannot use simple frequencies

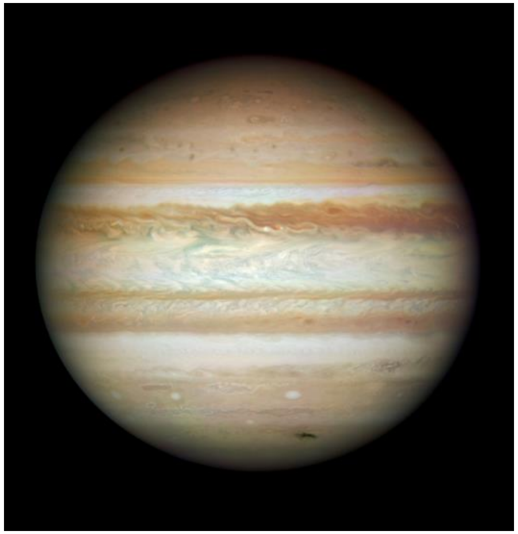


Image from https://www.wikimedia.org

From Sivia (1996), What is the mass of Jupiter?

Frequentist: measure the mass of enough Jupiter-like planets from multiple solar systems.

Bayesian: form a prior probability and update with any available information.

Probability Definitions
Bayesian Statistics

Bayes' Theorem:

Make an easy adjustment and we get the popular form.

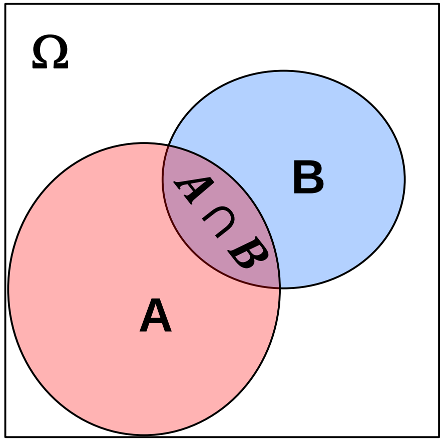
$$P(A|B) = \frac{P(B|A)\, P(A)}{P(B)}$$

Observations:

1. We are able to get P(A | B) from P(B | A) as you will see this often comes in handy.
2. Each term is known as:

$$\text{Posterior} = \frac{\text{Likelihood x Prior}}{\text{Evidence}}$$

3. Prior should have no information from likelihood.
4. Evidence term is usually just a standardization to ensure closure.



Venn Diagram – illustrating intersection.

Note: we got to Bayes' Theorem by fundamental frequentist approaches.

Probability Definitions
Bayesian Statistics

Bayes' Theorem:

A common example of Bayes' Theorem for modeling the subsurface.

Model Updating with a New Data Source:

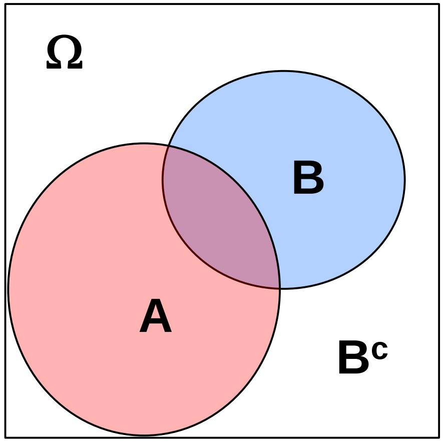
$$\underbrace{P(\text{Model} \mid \text{New Data})}_{\textbf{Posterior}} = \frac{\overbrace{P(\text{New Data} \mid \text{Model})}^{\textbf{Likelihood}}\ \overbrace{P(\text{Model})}^{\textbf{Prior}}}{\underbrace{P(\text{New Data})}_{\textbf{Evidence}}}$$

Probability Definitions Bayesian Statistics

Bayes' Theorem:

Alternative form, symmetry:

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



Venn Diagram – illustrating intersection.

Alternative form to calculate evidence term:

Given: $P(A) = \underbrace{P(A|B)\,P(B)}_{P(A \text{ and } B)} + \underbrace{P(A|B^c)\,P(B^c)}_{P(A \text{ and } B^c)}$

$$P(B|A) = \frac{P(A|B)\,P(B)}{P(A)} = \frac{P(A|B)\,P(B)}{P(A|B)\,P(B) + P(A|B^c)\,P(B^c)}$$

Probability Definitions
Bayesian Statistics

Bayesian approaches allow you to solve statistical problems that are not possible otherwise. Consider:

Event A	Event B
You have a disease	You test positive for the disease
There is fault compartmentalization	Geologist says there’s a fault
Low permeability of a sample	The laboratory measure is low
A valve will fail	X-ray test is positive
You drill a dry well	Seismic AVO response looks poor

In all of these cases you need to calculate:

$$P\left(\begin{matrix}\text{Something is}\\ \text{Happening}\end{matrix} \;\middle|\; \begin{matrix}\text{Looks like}\\ \text{its happening}\end{matrix}\right) = \frac{P\left(\begin{matrix}\text{Looks like}\\ \text{its happening}\end{matrix} \;\middle|\; \begin{matrix}\text{Something is}\\ \text{Happening}\end{matrix}\right) P\left(\begin{matrix}\text{Something is}\\ \text{Happening}\end{matrix}\right)}{P\left(\begin{matrix}\text{Looks like}\\ \text{its happening}\end{matrix}\right)}$$

Probability Definitions
Bayesian Statistics Example

Bayesian approaches allow you to solve statistical problems that are not possible otherwise. Consider:

Correct Detection Rate x Occurrence Rate

$$P\left(\begin{matrix}\text{Something is}\\\text{Happening}\end{matrix} \;\middle|\; \begin{matrix}\text{Looks like}\\\text{its happening}\end{matrix}\right) = \frac{P\left(\begin{matrix}\text{Looks like}\\\text{its happening}\end{matrix} \;\middle|\; \begin{matrix}\text{Something is}\\\text{Happening}\end{matrix}\right) P\left(\begin{matrix}\text{Something is}\\\text{Happening}\end{matrix}\right)}{P\left(\begin{matrix}\text{Looks like}\\\text{its happening}\end{matrix}\right)}$$

All Detection Rate (included false positives)

Often these terms are much easier to collect:

$$P(B|A) = \frac{P(A|B)\,P(B)}{P(A)} = \frac{P(A|B)\,P(B)}{P(A|B)\,P(B) + P(A|B^c)\,P(B^c)}$$

Let's try this out next.

Probability Definitions
Bayesian Statistics Hands-on

Example: Prior information at a site suggests a channel feature exists at a given location with probability of 60%. We decide to further investigate this information using seismic data.

$P(A) = 0.6$

$P(B|A) = 0.9$

$P(B^c|A^c) = 0.7$

$P(A^c) =$

$P(B|A^c) =$

Seismic survey can show a feature
- is present with 90% probability if it really is present
- is not present with a probability 70% if it really is not

***A*=The feature is present**
***B* =Seismic shows the feature**
***A^c* =The feature not present**
***B^c* =Seismic does not show the feature**

Will seismic information be useful?

Recall:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B|A)\,P(A) + P(B|A^c)\,P(A^c)}$$

Probability Definitions
Bayesian Statistics Example

Example: Prior information at a site suggests a channel feature exists at a given location with probability of 60%. We decide to further investigate this information using seismic data.

Seismic survey can show a feature
- is present with 90% probability if it really is present
- is not present with a probability 70% if it really is not

$P(A) = 0.6$
$P(B|A) = 0.9$
$P(B^c|A^c) = 0.7$
$P(A^c) = 1 - P(A) = 0.4$
$P(B|A^c) = 1 - P(B^c|A^c) = 0.3$

***A*=The feature is present**
***B* =Seismic shows the feature**
***A^c* =The feature not present**
***B^c* =Seismic does not show the feature**

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)} = \frac{P(B|A)\,P(A)}{P(B|A)\,P(A) + P(B|A^c)\,P(A^c)} = \frac{\overbrace{(0.9)(0.6)}^{\text{True Positive}}}{\underbrace{(0.9)(0.6)}_{\text{True Positive}} + \underbrace{(0.3)(0.4)}_{\text{False Positive}}} = 82\%$$

Will seismic information be useful? $P(A) = 60\%$ and now $P(A|B) = 82\%$

Bayesian Hands-on
Actual Happening Given Indicator

Bayesian Updating V2.0 - Inverting Conditional Probabilities

Michael Pyrcz, the University of Texas at Austin, Geostatistical Reservoir Modeling Class, @GeostatsGuy

With **Bayesian Updating** we can invert conditional probabilities (e.g. ***P(A|B)* → *P(B|A)***. This is very powerful, because often we can use an easier to calculate conditional probability to assess a more difficult to calculate, but more important conditional probability. For example, your doctor gives you a medical test that comes back positive for a disease. It would be important to know what is the probability that you have the disease given the positive test. This is a general category of problems that may be generalized as follows. **You have an positive indicator that something is happening, is the thing actually happening?** E.g. seismic interpretation indicates a fault, x-ray analysis indicates a crack etc.

$$P(\text{Actually Happening} \mid \text{Positive Indicator}) = \frac{P(\text{Positive Indicator} \mid \text{Actually Happening}) \times P(\text{Actually Happening})}{P(\text{Positive Indicator})}$$

It turns out that the denominator (Evidence Term) is often hard to calculate so we may use probability logic to calculate it as follows:

$$P(\text{Positive Indicator}) = \underbrace{P(\text{Positive Indicator} \mid \text{Actually Happening}) \times P(\text{Actually Happening})}_{\text{True Positive}} + \underbrace{P(\text{Positive Indicator} \mid \text{NOT Actually Happening}) \times P(\text{NOT Actually Happening})}_{\text{False Positive}}$$

Returning to the doctor's office. Your doctor has just informed you that you have tested positive (Positive Indicator) for a disease. Don't panic, resort to probability math. What infromation do you have to work with?

Instructions:
Adjust the yellow probabilities (that would likely be available) and observe the resulting probability of having the disease given a positive test. Note intermediate calculated probabilities are in blue cells.

Probability of getting this disease	***P(Actually Happening)*** =	0.001%
By closure the compliment, probability of not getting this disease	***P(Not Actually Happening)* = *1 – P(Actually Happening)*** =	99.999%
Probability of detecting the disease if you have it. This is the sensitivity of the test.	***P(Positive Indicator / Actually Happening)*** =	99.000%
Probability of detecting the disease if you don't have it. This is the false positive rate of the test.	***P(Positive Indicator / NOT Actually Happening)*** =	0.010%

P(Positive Indicator) = $P(\text{Positive Indicator} \mid \text{Actually Happening}) \times P(\text{Actually Happening}) + P(\text{Positive Indicator} \mid \text{NOT Actually Happening}) \times P(\text{NOT Actually Happening})$

P(Positive Indicat 0.99% x 0.00001% + 0.0001% x 0.99999% → ***P(Positive Indicator)*** = 0.011%

We now have everything we need to solve for the probability you have the disease given a postive test.

$$P(\text{Actually Happening} \mid \text{Positive Indicator}) = \frac{P(\text{Positive Indicator} \mid \text{Actually Happening}) \times P(\text{Actually Happening})}{P(\text{Positive Indicator})}$$

$$= \frac{99.000\% \times 0.001\%}{0.011\%} =$$ ***P(Actually Happening / Positive Indicator)*** = **9.008%**

What should you observe?

Why is the ***P(Actually Happening / Positive Indicator)*** so low? Check out the following joint probabilities.

The probability of experiencing a false positive is ***P(Not Actually Happening and Positive Indicator)*** = 0.010% | 10.10 Ratio of Probability of False Positive / Probability of True Positive

Compare this to the true positive ***P(Actually Happening and Positive Indicator)*** = 0.001%

The combination of a very unlikely event (rare disease) and a signigicant false positive rate results in 10.1x greater probability of a false positive than a true positive with this test. The problem is that given an apparently low false positive rate and a very high true positive rate most people would assume that the detected condition is actually happening, when in fact it is unlikely!

For more (geo)statistical demos check out github/GeostatsGuy and twitter @GeostatsGuy.

False Positives:
What if false positive dropped form 0.01% to 0.001%?

Rare Events:
What if probability of occurance increased from 0.001% to 0.01%?

Example at https://github.com/GeostatsGuy/ExcelNumericalDemos/blob/master/BayesianUpdatingInversion_Demo.xlsx

Probability Definitions
Bayesian Statistics Example

Machine 1

$P(X_1)$, **20% Production**
$P(Y|X_1)$, **5% Error Rate**

Machine 2

$P(X_2)$, **30% Production**
$P(Y|X_2)$, **3% Error Rate**

Machine 3

$P(X_3)$, **50% Production**
$P(Y|X_3)$, **1% Error Rate**

Example: Probability of an error in the product, $P(Y)$?

Hint: Calculate Marginal $P(Y) = \sum_{i=1}^{n} P(Y, X_i) = \sum_{i=1}^{n} P(Y|X_i)P(X_i)$

Example https://en.wikipedia.org/wiki/Bayes%27_theorem

Probability Definitions
Bayesian Statistics Example

Machine 1

$P(X_1)$, **20% Production**
$P(Y|X_1)$, **5% Error Rate**

Machine 2

$P(X_2)$, **30% Production**
$P(Y|X_2)$, **3% Error Rate**

Machine 3

$P(X_3)$, **50% Production**
$P(Y|X_3)$, **1% Error Rate**

Example: Probability of an error in the product, $P(Y)$?

$$P(Y) = \sum_{i=1}^{n} P(Y, X_i) = \sum_{i=1}^{n} P(Y|X_i)P(X_i)$$

$$P(Y) = P(Y|X_1)P(X_1) + P(Y|X_2)P(X_2) + P(Y|X_3)P(X_3)$$

$$P(Y) = (0.20)(0.05) + (0.30)(0.03) + (0.50)(0.01)$$

$$P(Y) = 0.024 = 2.4\%$$

Example https://en.wikipedia.org/wiki/Bayes%27_theorem

Probability Definitions
Bayesian Statistics Example

Machine 1	Machine 2	Machine 3
$P(X_1)$, **20% Production**	$P(X_2)$, **30% Production**	$P(X_3)$, **50% Production**
$P(Y\|X_1)$, **5% Error Rate**	$P(Y\|X_2)$, **3% Error Rate**	$P(Y\|X_3)$, **1% Error Rate**

Example: Probability product came each machine given an error is observed, $P(X_i|Y)$?

Note: From the previous slide: $P(Y) = 0.024 = 2.4\%$

Hint calculate the conditional: $P(X_i|Y) = \frac{P(Y|X_i)P(X_i)}{P(Y)}$

Probability Definitions

Bayesian Statistics Example

Machine 1

$P(X_1)$, **20% Production**
$P(Y|X_1)$, **5% Error Rate**

Machine 2

$P(X_2)$, **30% Production**
$P(Y|X_2)$, **3% Error Rate**

Machine 3

$P(X_3)$, **50% Production**
$P(Y|X_3)$, **1% Error Rate**

Example: Probability product came each machine given an error is observed, $P(X_i|Y)$?

$$P(X_i|Y) = \frac{P(Y|X_i)P(X_i)}{P(Y)}$$

$$P(X_1|Y) = \frac{(0.05)(0.2)}{(0.024)} = 0.41$$

$$P(X_3|Y) = \frac{(0.01)(0.5)}{(0.024)} = 0.21$$

$$P(X_2|Y) = \frac{(0.03)(0.3)}{(0.024)} = 0.38$$

Example https://en.wikipedia.org/wiki/Bayes%27_theorem

Bayesian Hands-on
Updating Exploration Success

Induction with Bayes' Theorem for Updating Exploration Success Rate with Exploration Drilling Results

Michael Pyrcz, Associate Professor, the University of Texas at Austin

Problem: update the assumed exploration success rate with new exploration drilling results. Update the prior exploration probability of exploration success with n_s drilling successes out of n new exploration wells.

$$Prob\{\mathbf{Model}|Result\} = \frac{Prob\{\mathbf{Result}|\mathbf{Model}\} \cdot Prob\{Model\}}{Prob\{\mathbf{Result}\}}$$

we can use Bayes' Theorem go from Prob{Result | Model}(probability of exploration drilling outcome given exploration model) that is easy to calculate to the Prob{Model | Outcome}(probabilty of the exploration model success rate given drilling outcomes) that is not available.

$$Posterior = \frac{Likelihood \cdot \mathbf{Prior}}{Evidence}$$

the prior is our belief of the probability of each possible exploration success rate (an uniform probability distribution is a naïve prior - we don't know) before drilling the new exploration wells.
Likelihood comes from the binomial distribution. Evidence is the normalization constant such that the resulting posteriori PDF sums to 1.0.

$$Prob\{Result\} = \mathbf{k} \qquad Prob\{\mathbf{Result}|\mathbf{Model}\} = \binom{n_s}{n} P(n_s)^{n_s} \cdot (1 - P(n_s))^{n-n_s}, n_s$$

where n_s is the number successes, n is the total number of wells and $P(n_s)$ is the probability of exploration success.

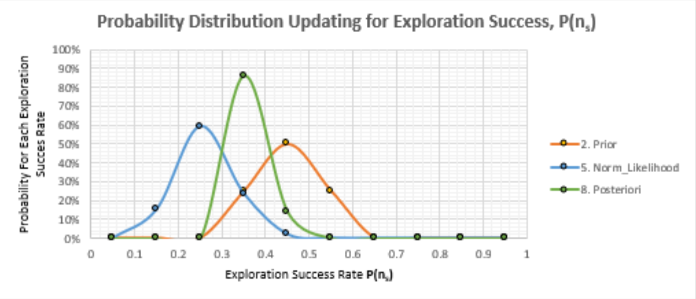
1. Data results - Exploration Outcome

Heads, n_s	10
Failures, $n - n_s$	30

Experimental Exploration Success Rate: 33.3%

Prior, Likelihood and Posterior probabilities Binned by Probability of Exploration Drilling Success

Prob. of Heads, H	0.05	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95	Sum
2. Prior	0.00000	0.00000	0.00000	0.25000	0.50000	0.25000	0.00000	0.00000	0.00000	0.00000	1.0000
3. Norm_Prior	0.00000	0.00000	0.00000	0.25000	0.50000	0.25000	0.00000	0.00000	0.00000	0.00000	
4. Likelihood	0.00002	0.03730	0.14436	0.05706	0.00469	0.00008	0.00000	0.00000	0.00000	0.00000	0.2435
5. Norm_Likelihood	0.00007	0.15317	0.59284	0.23430	0.01926	0.00035	0.00000	0.00000	0.00000	0.00000	
6. Prior x Likelihood	0.00000	0.00000	0.00000	0.01426	0.00235	0.00002	0.00000	0.00000	0.00000	0.00000	0.0166
7. Evidence	0.01663	0.01663	0.01663	0.01663	0.01663	0.01663	0.01663	0.01663	0.01663	0.01663	
8. Posteriori	0.00000	0.00000	0.00000	0.85770	0.14102	0.00127	0.00000	0.00000	0.00000	0.00000	1.0000



Instructions for Bayes' Theorem Excel Demo

1. Set any data outcome, Data. Where Heads, n_s, is the number of exploration sucesses and $n - n_s$ is the number of epxloration failures and n is the total number of exploration wells.
2. Set the prior to any set of relative probabilities to reflect prior belief concerning the exploration drilling success rate prior to drilling the new exploration wells. Constant is a naïve prior (no idea) or higher for 0.4 reflects a prior / belief in a 40% exploration success rate.
3. The prior probabilities for each exploration success rate bin are standardized to sum to 1.0 as expected for a PDF.
4. The likelihood calculated from the binomial distribution based on the exploration drilling outcome.
5. The likelihood normalized sum to 1.0 as expected for a PDF (for plotting).
6. The product of the prior and the likelihood.
7. The evidence term as the sum of the product of prior and likelihood to ensure the posteriori sums to 1.0 over the exploration success rate bins as expected for a PDF.
8. The posterior as the product of prior and likelihood standardized by evidence for each exploration success rate bin.

What did we learn?

1. Bayes' Theorem may be applied to calculated conditional probabilities that otherwise would be difficult to assess.
2. The prior model has a significant impact on the posterior and must be selected carefully.
3. For a naïve prior the posterior is equal to the likelihood.

Based on Sivia, D.S., 1996, Data Analysis, A Bayesian Tutorial, Oxford Science Publications, 189 p.

Example at https://github.com/GeostatsGuy/ExcelNumericalDemos/blob/master/Bayesian_Exploration_Demo.xlsx

Bayesian Updating with Gaussian Distributions

There is an analytical solution for working with Gaussian distributions for Bayesian updating (Sivia, 1996).

- Calculate the mean of the posterior from the prior and likelihood mean and variance.

$$\bar{x}_{\text{updated}} = \frac{\bar{x}_{\text{likelihood}}(\mathbf{u}) \cdot \sigma^2_{\text{prior}}(\mathbf{u}) + \bar{x}_{\text{prior}}(\mathbf{u}) \cdot \sigma^2_{\text{likelihood}}(\mathbf{u})}{[1 - \sigma^2_{\text{likelihood}}(\mathbf{u})][\sigma^2_{\text{prior}}(\mathbf{u}) - 1] + 1}$$

- Calculate the variance of the posterior form the prior and likelihood variances (no means, homoscedastic!).

$$\sigma^2_{\text{updated}}(\mathbf{u}) = \frac{\sigma^2{}_{\text{prior}}(\mathbf{u})\,\sigma^2{}_{\text{likelihood}}(\mathbf{u})}{[1 - \sigma^2{}_{\text{likelihood}}(\mathbf{u})][\sigma^2{}_{\text{prior}}(\mathbf{u}) - 1] + 1}$$

Bayesian Hands-on
Updating with Gaussian Distributions

Bayesian Gaussian Analytical Example Demo

Michael Pyrcz, the University of Texas at Austin, Geostatistical Reservoir Modeling Class

Formulation from Sivia, 1996.

$$\bar{x}_{\text{updated}} = \frac{\bar{x}_{\text{likelihood}}(\mathbf{u}) \cdot \sigma^2_{\text{prior}}(\mathbf{u}) + \bar{x}_{\text{prior}}(\mathbf{u}) \cdot \sigma^2_{\text{likelihood}}(\mathbf{u})}{[1 - \sigma^2_{\text{likelihood}}(\mathbf{u})][\sigma^2_{\text{prior}}(\mathbf{u}) - 1] + 1}$$

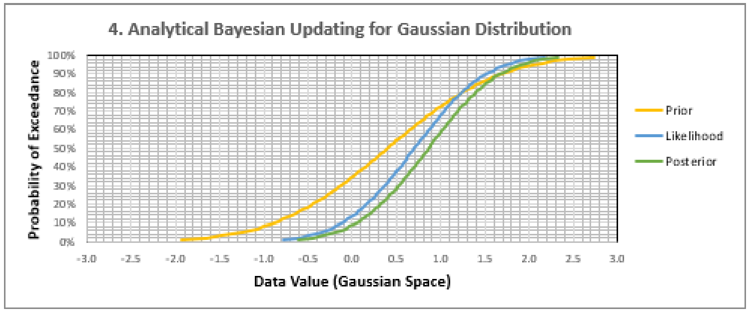
$$\sigma^2_{\text{updated}}(\mathbf{u}) = \frac{\sigma^2_{\text{prior}}(\mathbf{u})\,\sigma^2_{\text{likelihood}}(\mathbf{u})}{[1 - \sigma^2_{\text{likelihood}}(\mathbf{u})][\sigma^2_{\text{prior}}(\mathbf{u}) - 1] + 1}$$

1. Prior Distribution	
average	0.40
variance	1.00

2. Likelihood Distribution	
average	0.70
variance	0.40

3. Posterior Distribution	
average	0.86
variance	0.40

Percentile	Prior	Likelihood	Posterior
0.01	-1.926	-0.771	-0.611
0.02	-1.654	-0.599	-0.439
0.03	-1.481	-0.490	-0.330
0.04	-1.351	-0.407	-0.247
0.05	-1.245	-0.340	-0.180
0.06	-1.155	-0.283	-0.123
0.07	-1.076	-0.233	-0.073
0.08	-1.005	-0.189	-0.029
0.09	-0.941	-0.148	0.012
0.1	-0.882	-0.111	0.049
0.11	-0.827	-0.076	0.084
0.12	-0.775	-0.043	0.117
0.13	-0.726	-0.012	0.148
0.14	-0.680	0.017	0.177
0.15	-0.636	0.045	0.205
0.16	-0.594	0.071	0.231
0.17	-0.554	0.097	0.257
0.18	-0.515	0.121	0.281
0.19	-0.478	0.145	0.305
0.2	-0.442	0.168	0.328
0.21	-0.406	0.190	0.350
0.22	-0.372	0.212	0.372
0.23	-0.339	0.233	0.393
0.24	-0.306	0.253	0.413
0.25	-0.274	0.273	0.433
0.26	-0.243	0.293	0.453
0.27	-0.213	0.312	0.472
0.28	-0.183	0.331	0.491
0.29	-0.153	0.350	0.510
0.3	-0.124	0.368	0.528
0.31	-0.096	0.386	0.546
0.32	-0.068	0.404	0.564
0.33	-0.040	0.422	0.582
0.34	-0.012	0.439	0.599
0.35	0.015	0.456	0.616
0.36	0.042	0.473	0.633



Instructions for Analytical Bayesian Updating for Gaussian Distributions

1. Set the average and the variance of the prior distribution (Gaussian parametric distribution).
2. Set the average and the variance of the likelihood distribution (Gaussian parametric distribution).
3. Observed the updated average and variance of the posterior distribution (Gaussian parametric distribution).
4. Observed the prior, likelihoof and posterior cumulative distribution functions (CDFs).

What did we learn?

1. The posterior variance is only a function of the prior and likelihood variances. The prior and likelihood means have no influence.
2. In general updating results in a reduction variance. Posterior variance is equal to or less than the greater of the prior and the likelihood variance.
3. High certainty in either prior or likelihood distribution (very low variance) causes either term to dominate the updated posterior.

Sivia, D.S., 1996, Data Analysis, A Bayesian Tutorial, Oxford Science Publications, 189 p.

Example https://github.com/GeostatsGuy/ExcelNumericalDemos/blob/master/Bayesian_Gaussian_Demo.xlsx

Probability New Tools

Topic	Application to Subsurface Modeling
Frequentist Concepts	When sufficient observations are available use (long-run) counting to access the required probabilities. *Predict reservoir average porosity by pooling analogous fields.*
Bayesian Concepts Inversion of Conditionals	Calculate a difficult to access conditional probability from an accessible one. Probability of event given indicator from probability indicator given event. *Calculate probability of sealing fault given indicator of sealing fault.*
Bayesian Concepts Bayesian Updating	Update prior belief with new information. *Calculate probability of exploration success rate given prior model and outcomes from exploration drilling program.*

Multivariate Modeling: Probability and Statistics

Lecture outline . . .

- Probability in Subsurface Modeling
- Frequentist Concepts
- Bayesian Concepts

Introduction

Fundamental Concepts

Probability

Data Prep / Analytics

Spatial Continuity / Prediction

Multivariate Modeling

Uncertainty Modeling

Machine Learning

Instructor: Michael Pyrcz, the University of Texas at Austin