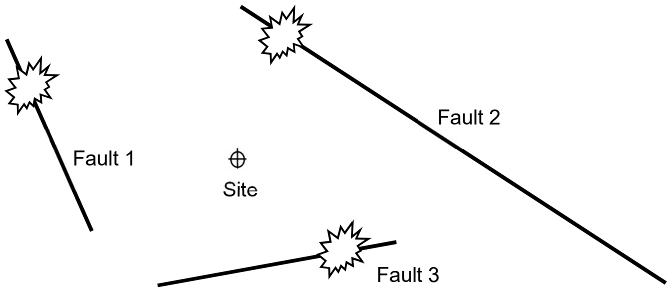
Instructor: Junho Song
junhosong@snu.ac.kr

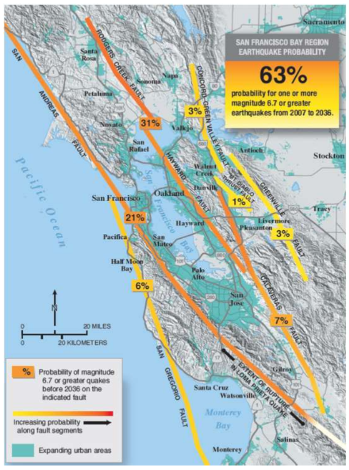
457.212 Statistics for Civil & Environmental Engineers
In-Class Material: Class 08
Total Probability Theorem & Bayes Rule (A&T: 2.3)

1. **Total probability theorem:** “   ” and “   “ approach

   (a) Example: Under seismic hazard from multiple faults surrounding a building site, what is the probability of the event E: the peak ground acceleration of a earthquake is greater than 0.3g, i.e. $PGA > 0.3$?

   $P(E)$ is needed to determine the appropriate level of seismic design (new) or retrofit (existing).

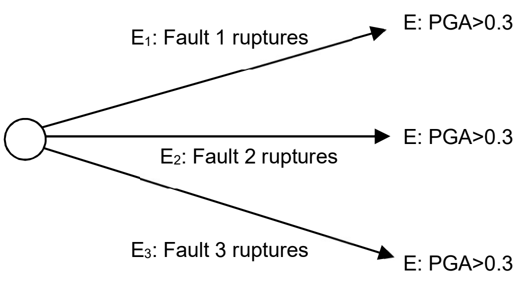


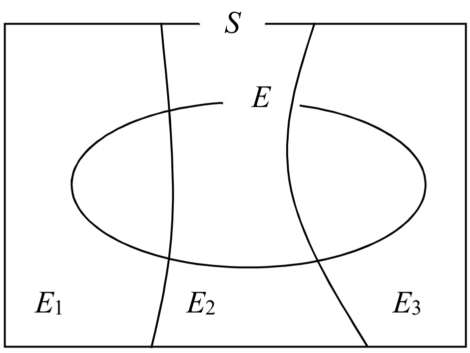


https://www.shakeout.org/california/bayarea/

Issue: each fault has different distance from the site, length,
depth, geological characteristics although they affect the intensity of ground motion significantly.

Solution: **tackle one by one.**





(b) **Theorem:** If E_i's , $i = 1,\ldots,n,$ are (   ) and (   ),

$$
\begin{aligned}
P(E) &= P(\quad \cup \quad \cup \cdots \cup \quad) \\
&= P(\quad) + P(\quad) + \cdots + P(\quad) \\
&= P(\quad | \quad)P(E_1) + \cdots P(\quad | \quad)P(E_n) \\
&= \sum_{i=1}^{n} P(E \mid E_i)P(E_i)
\end{aligned}
$$

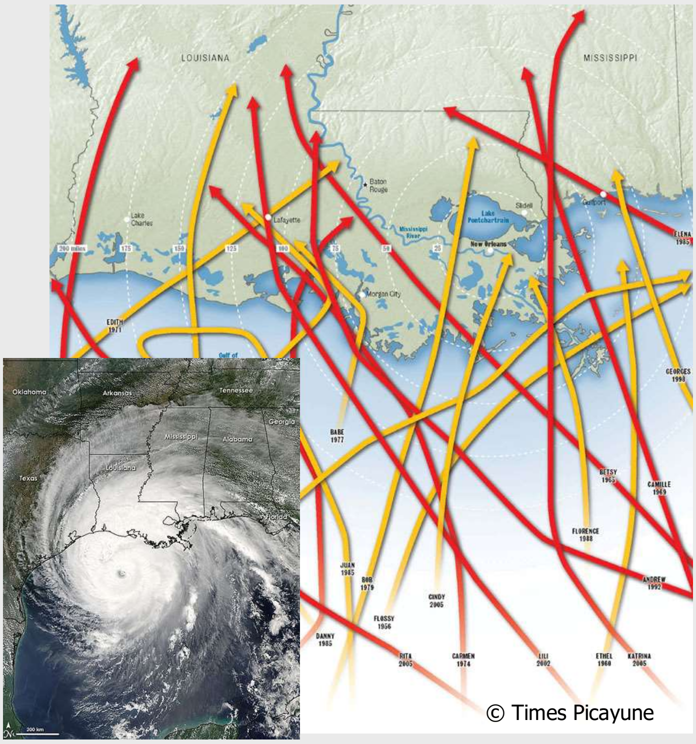
Example 1 (A&T 2.25): Consider landfalls of hurricanes at a particular area in the southern coast of Louisiana (assume: at most one hurricane landfall per year)

Saffir/Simpson Hurricane Category	C1	C2	C3	C4	C5
Velocity (mph)	74-95	96-110	111-130	131-155	≥156
Probability (/yr)	0.35	0.25	0.14	0.05	0.01
Conditional Prob.of Damage, $P(D\mid C_i)$	0.05	0.10	0.25	0.60	1.00

C0: Velocity < 74 mph
P(C0)?

Note P(D|C0) = 0

The annual probability of wind damage of the building, P(D)?



2. **Bayes theorem:** $P(E_i \mid E)$ from $P(E \mid E_i)$'s and $P(E_i)$'s – an inverse formula

$$P(E_i \mid E) = \frac{P(E \mid E_i)}{P(E)} \cdot P(E_i)$$

(a) Proof?

(b) $P(E)$?

(c) Bayes theorem is everywhere – pattern recognition (face, fingerprint, hand-writing, customer behavior), model parameter estimation, decision-making, etc.

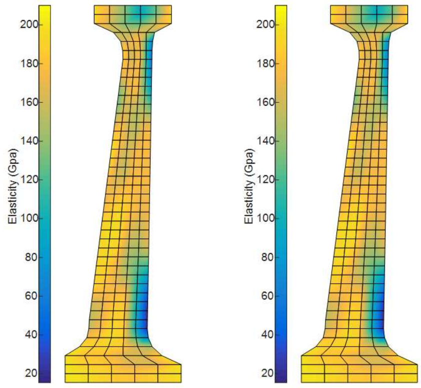
(d) Represents a philosophy or definition of probability: ________ of _________

(d) CEE Example 1: Identification of structural damage from indirect measurement data

> Lee, S.-H., and J. Song (2017). System identification of spatial distribution of structural parameters using modified Transitional Markov Chain Monte Carlo (m-TMCMC) method. *ASCE Journal of Engineering Mechanics*. Vol. 143(9), 04017099-1~18.

E_i: Damage pattern
E: Measured displacement under loads

Method: modified version of transitional Markov Chain Monte Carlo method (m-TMCMC)



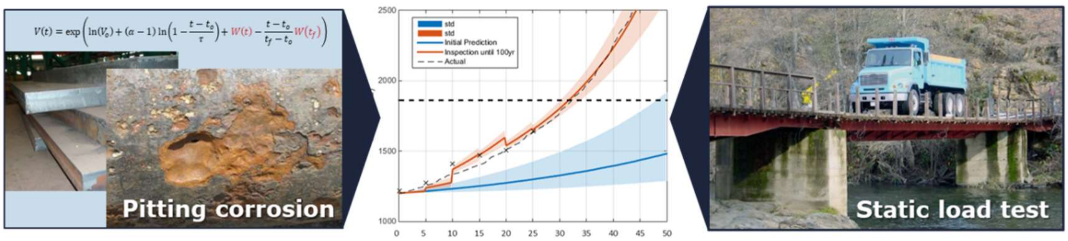
(a) Reference scenario (b) Bayesian estimation

(e) CEE Example 2: Monitoring and prediction of deterioration

> Yi, S., and J. Song. Particle filter based monitoring and prediction of spatiotemporal deterioration using successive measurements of structural responses. *Mechanical System and Signal Processing*, Under Review.

E_i: Deterioration model parameter
E: Measured displacement/strain under loads

Method: Bayesian filters, e.g. particle filter



Corrosion progress model (stochastic model) | Bayesian filter (particle filter) | Indirect measurements (strain with noise)

Instructor: Junho Song
junhosong@snu.ac.kr

Example 2 (A&T 2.29): Aggregates are supplied by two companies (A & B) for the construction of a reinforced concrete (RC) building.

- Company A: 600 truck loads/day, 3% disqualification rate
- Company B: 400 truck loads/day, 1% disqualification rate

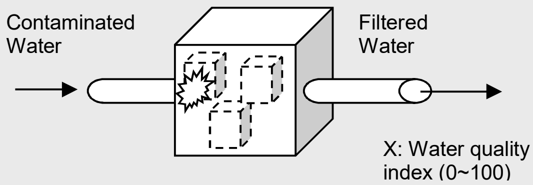
Consider the events

- A: a load of aggregate picked at random came from Company A
- B: a load of aggregate picked at random came from Company B
- E: a load of aggregate picked at random is substandard and thus disqualified

(a) $P(A)$ and $P(B)$

(b) If the load of aggregate picked at random is substandard, what is the probability that it came from Company A?

Example 3: Consider a water purification system consisting of three major components. Assume at most one component can be out of order at a time. Suppose they use a water quality index, X that ranges from 0 (clean) to 100 (contaminated).



Instructor: Junho Song
junhosong@snu.ac.kr

Suppose we know

Failed Component	$P(E_i)$	$P(0 \le X \le 20 \mid E_i)$	$P(70 \le X \le 100 \mid E_i)$
1st comp. (E_1)	0.3	0.9	0.05
2nd comp. (E_2)	0.2	0.3	0.4
3rd comp. (E_3)	0.1	0.1	0.8
No failures (E_0)		1.0	0

(a) $P(0 \le X \le 20)$

(b) $P(70 \le X \le 100)$

(c) If $0 \le X \le 20$, what is the probability that the *i*-th component has failed, $i = 0,1,2,3$?

(d) If $70 \le X \le 100$, what is the probability that the i-th component has failed, $i = 0,1,2,3$?

Failed Component		
1st comp. (E_1)		
2nd comp. (E_2)		
3rd comp. (E_3)		
No failures (E_0)		