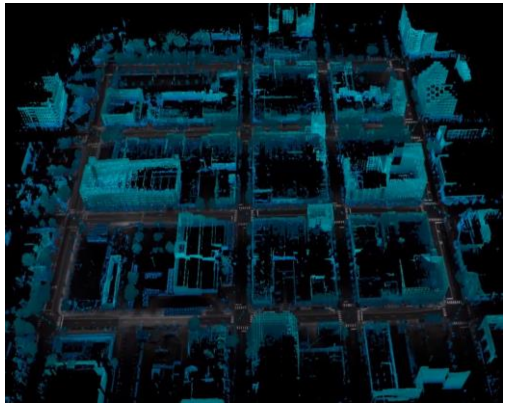
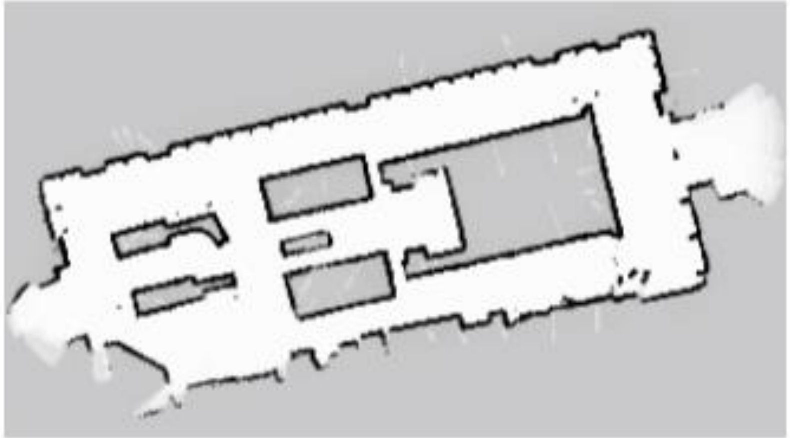
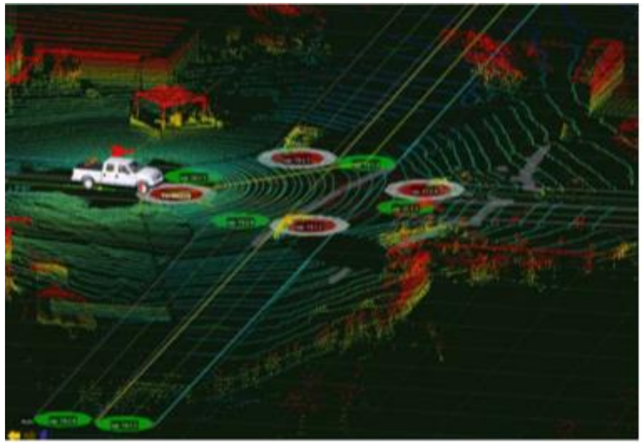
BYU Electrical & Computer Engineering





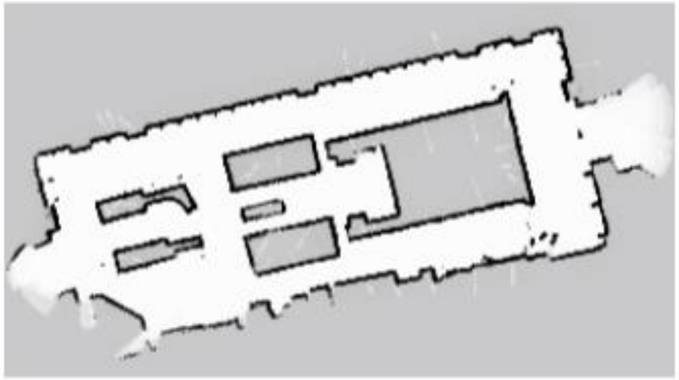
# PROBABILITY REVIEW - BASICS

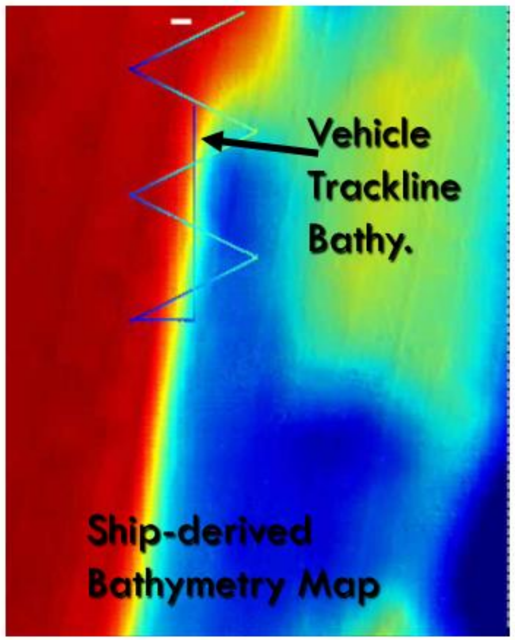
## ECEN 633: Robotic Localization and Mapping

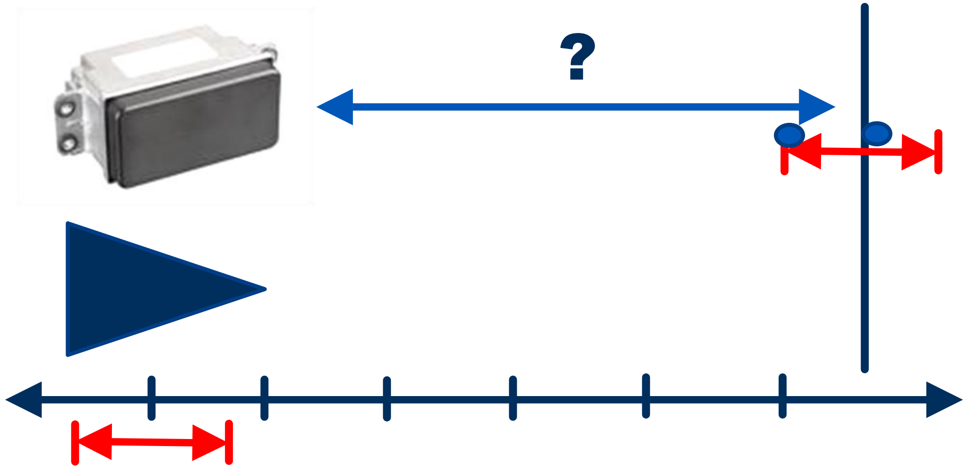
Slides Based on probabilistic-robotics.org and a slide deck by Ryan Eustice.

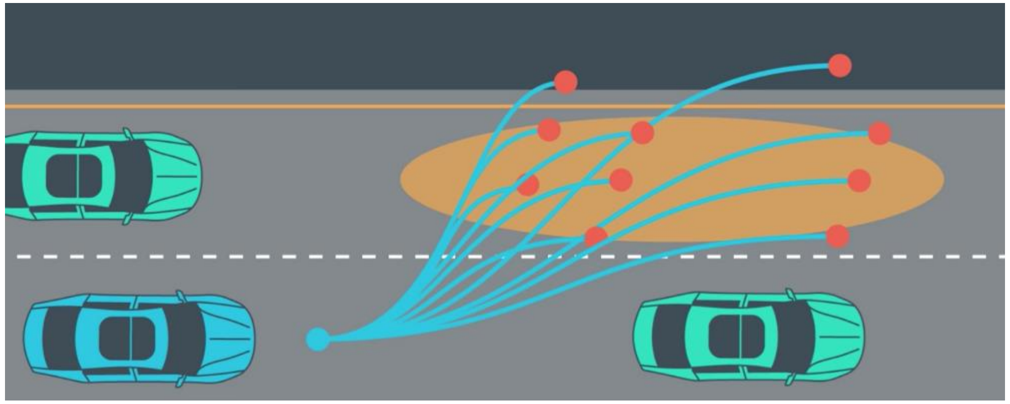
# Dealing With Uncertainty

- Fundamental Problems in Robotics:
  - Mapping
  - Localization
  - Perception
  - Planning





# The Axioms of Probability

$\Omega$ is the set of all possible outcomes of an experiment.
$\Pr(A)$ denotes the the probability that proposition $A \subset \Omega$ is true.

1. $0 \leq \Pr(A) \leq 1 \quad \forall$ valid $A \subset \Omega$

2. $\Pr(\Omega) = 1$

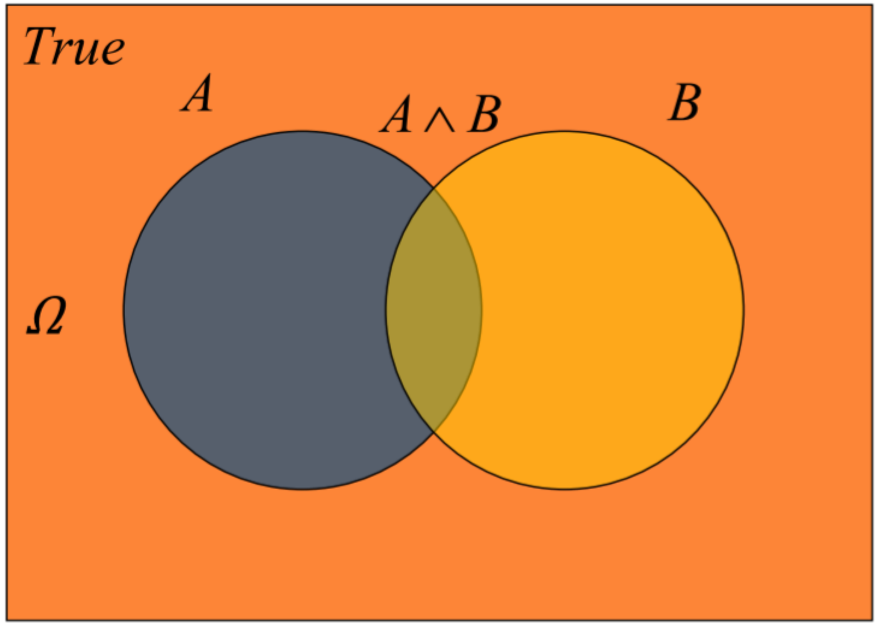
3. Any countable sequence of disjoint sets in $\Omega$ satisfies:
$$\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr\left(A_i\right)$$

# Axiom 3

Any countable sequence of disjoint sets in $\Omega$ satisfies:

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr\left(A_i\right)$$

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$



# Using the Axioms

$$
\begin{aligned}
\Pr(A \cup \neg A) = & \quad \Pr(A) + \Pr(\neg A) - \Pr(A \cap \neg A) \\
\Pr(\text{True}) = & \quad \Pr(A) + \Pr(\neg A) - \Pr(\text{False}) \\
1 & = \quad \Pr(A) + \Pr(\neg A) - 0 \\
\Pr(\neg A) & = \quad 1 - \Pr(A)
\end{aligned}
$$

# Random Variables

**DEF: A *Random Variable* is a function that maps the outcomes of a random experiment to a real number or a set of real numbers.**

$$\mathbf{X} : \mathbf{\Omega} \to \mathbb{R}$$

# Discrete Random Variables

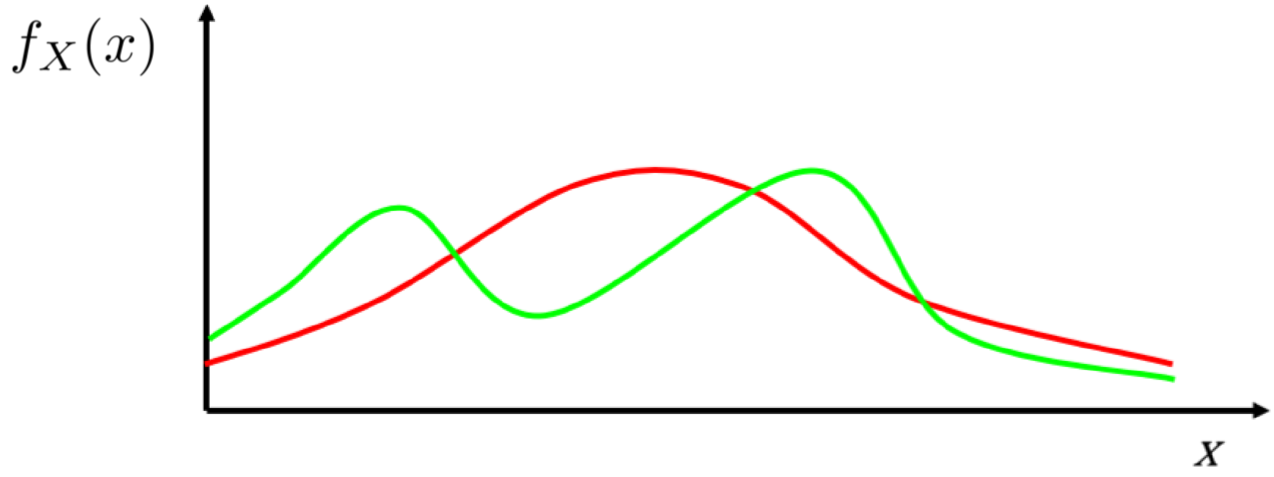
- Take on a discrete countable number of values $\{x_1, x_2, x_3, \cdots, x_n\}$
- $P(X = x_i)$ or $P(x_i)$ is the probability that the random variable $X$ takes on the value $x_i$
- $P(x) = f_X(x)$ is called a Probability Mass Function or PMF

# Continuous Random Variables

- Take on values in a continuous range

- $P(a \leq x \leq b)$ is the probability that the random variable takes on a value in the range from $a$ to $b$

$$P(a \leq x \leq b) = \int_a^b f_X(x)dx$$

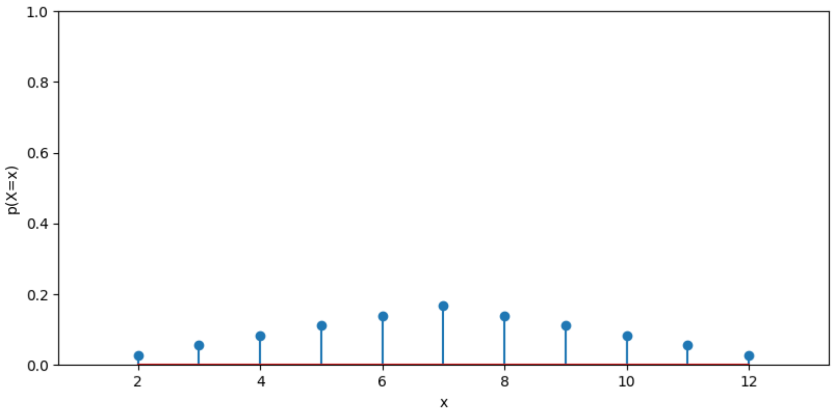



$f_X$ is called a Probability Density Function or PDF

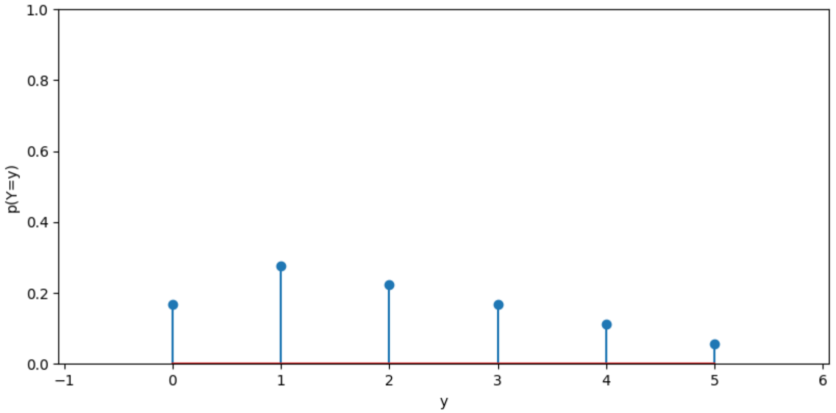
# Joint Probability Distribution

- Jointly describes the probability that multiple random variables take on specific values $P(x, y) = P(X = x \text{ and } Y = y)$

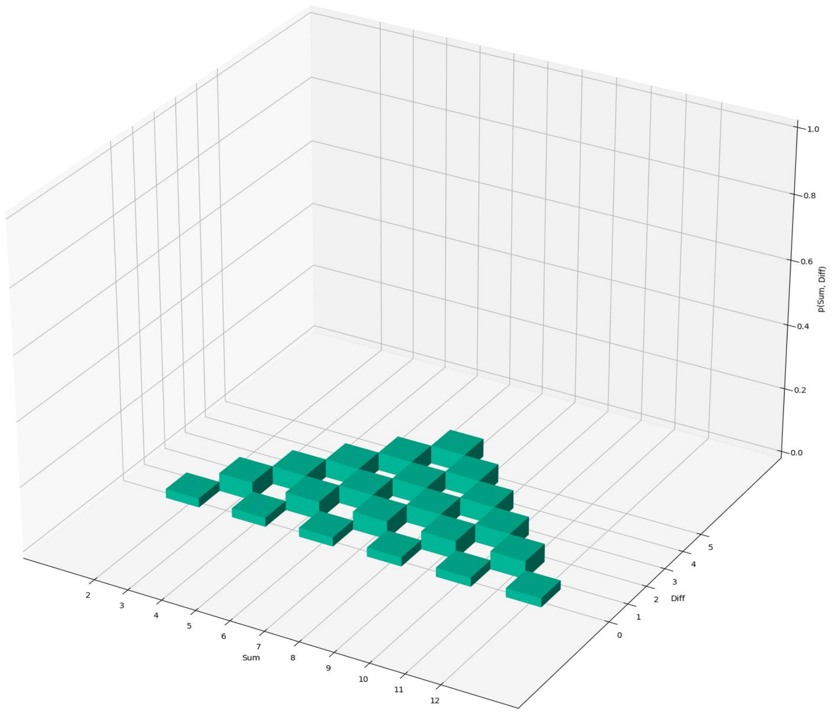
PMF of Sum of Two Dice




PMF of Diff. of Two Dice



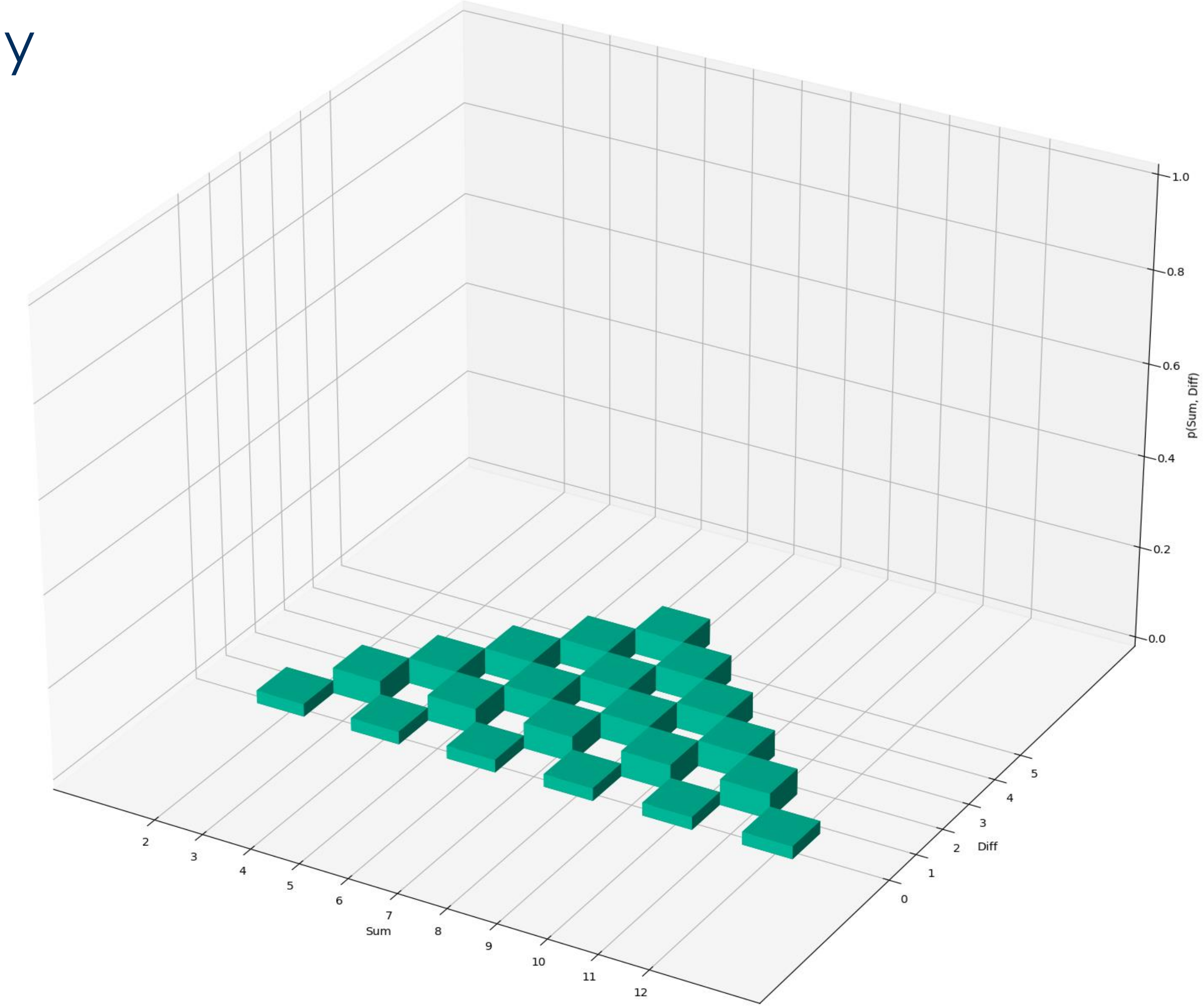

Joint PMF of Sum and Diff.



# Conditional Probability and Independence

- $P(x|y)$ is the probability of x *given* y

$$P(x \mid y) = \frac{P(x,y)}{P(y)}$$



# Conditional Probability and Independence

- X and Y are called *independent* if

$$P(x, y) = P(x)P(y)$$

# Conditional Probability and Independence

Def. of Conditional Probability

$$P(x \mid y) = \frac{P(x,y)}{P(y)}$$

$$P(y \mid x) = \frac{P(y,x)}{P(x)}$$

Def. of Independence

$$P(x, y) = P(x)P(y)$$

$\Rightarrow$

If X and Y are *independent*:

$$P(x \mid y) = P(x)$$

# Marginalization and the Law of Total Probability

| Discrete Case | Continuous Case |
| --- | --- |

**Axioms of Probability:**

$$\sum_x P(x) = 1 \qquad \int p(x)dx = 1$$

**Marginalization:**

$$P(x) = \sum_y P(x, y) \qquad p(x) = \int p(x, y)dy$$

**Law of Total Probability:**

$$P(x) = \sum_y P(x \mid y)P(y) \qquad p(x) = \int p(x \mid y)p(y)dy$$

# Bayes Formula

$$P(x,y) = P(x \mid y)P(y) = P(y \mid x)P(x)$$

$$\Rightarrow$$

$$P(x \mid y) = \frac{P(y \mid x)\, P(x)}{P(y)} = \frac{\text{likelihood} \cdot \text{prior}}{\text{evidence}}$$

$$P(x \mid y) = \frac{P(y \mid x)\, P(x)}{\sum_x P(y \mid x) P(x)}$$

via Law of Total Probability

# Normalization

$$P(x \mid y) = \frac{P(y \mid x)\ P(x)}{P(y)}$$

$$\eta = P(y)^{-1} = \frac{1}{\sum_x P(y \mid x) P(x)}$$

$$P(x \mid y) = \frac{P(y \mid x)\ P(x)}{P(y)} = \eta\ P(y \mid x)\, P(x)$$

# Checkpoint

| #Legs | Species | *P(L=#Legs, S=Species)* |
|---|---|---|
| 2 | Dog | 0.001 |
| 2 | Cat | 0.001 |
| 2 | Bird | 0.2 |
| 3 | Dog | 0.057 |
| 3 | Cat | 0.04 |
| 3 | Bird | 0.001 |
| 4 | Dog | 0.4 |
| 4 | Cat | 0.3 |
| 4 | Bird | 0 |

- P(#legs=2 U #legs=3 U #legs=4)
- P(Dog U Cat U Bird)
- P(Bird)
- P(Bird, #legs=2)
- P(Bird | #legs=2)
- P(#legs = 2 | Bird)

# Bayes Rule with Background Knowledge

$$P(x \mid y, z) = \frac{P(y \mid x, z)\; P(x \mid z)}{P(y \mid z)}$$

# Conditioning

- Law of Total Probability

$$p(x) = \int p(x,z)\,dz$$

$$p(x) = \int p(x \mid z)\,p(z)\,dz$$

$$p(x \mid y) = \int p(x \mid y,z)\,p(z \mid y)\,dz$$

# Conditional Independence

$$P(x, y \mid z) = P(x \mid z) P(y \mid z)$$

equivalent to

$$P(x \mid z) = P(x \mid z, y)$$

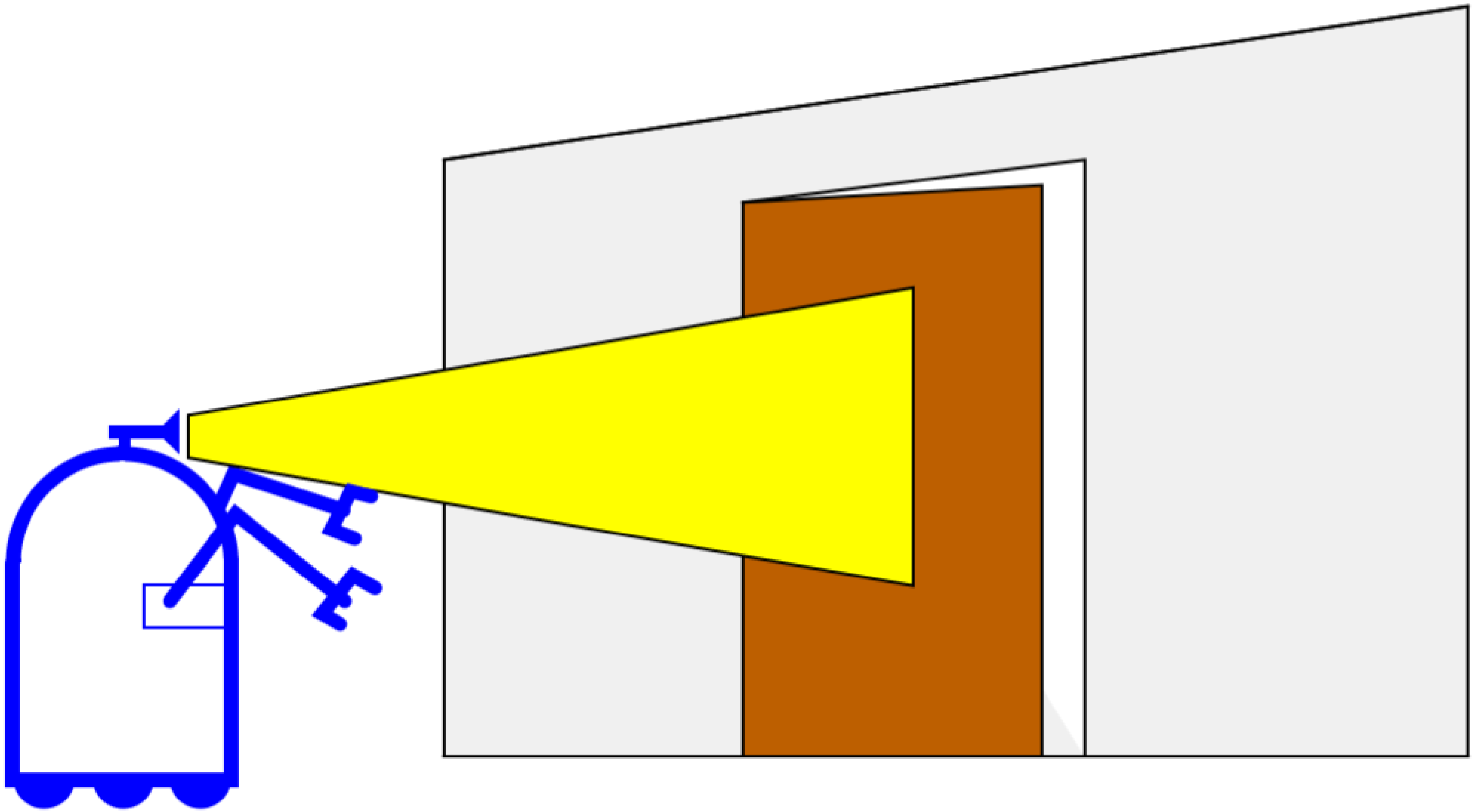
and

$$P(y \mid z) = P(y \mid z, x)$$

Conditional Independence does **NOT** imply Independence and Vice Versa!

# Simple State Estimation Example

- Suppose a robot obtains measurement z
  - E.g. robot estimates state of door using its camera
- What is P(open | z)



# Causal vs. Diagnostic Reasoning

- P( open | z ) is ***diagnostic***
- P( z | open ) is ***causal***
- Often causal knowledge is easier to obtain.
- Bayes rule allows us to swap them out

$$P(open \mid z) = \frac{P(z \mid open)P(open)}{P(z)}$$

# Bayes Rule Example

- $z = sense\_open$
- $P(z = sense\_open \mid open) = 0.6 \qquad P(z = sense\_open \mid \neg open) = 0.3$
- $P(open) = P(\neg open) = 0.5$

$$P(open \mid z) = \frac{P(z \mid open)P(open)}{P(z \mid open)P(open) + P(z \mid \neg open)P(\neg open)}$$

$$P(open \mid z = sense\_open) = \frac{0.6 \cdot 0.5}{0.6 \cdot 0.5 + 0.3 \cdot 0.5} = \frac{2}{3} = 0.67$$

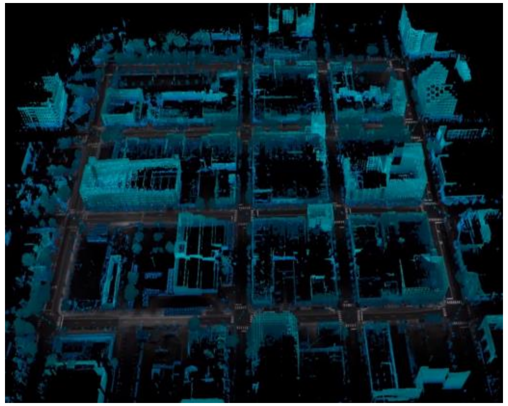
- z raises the probability that the door is open

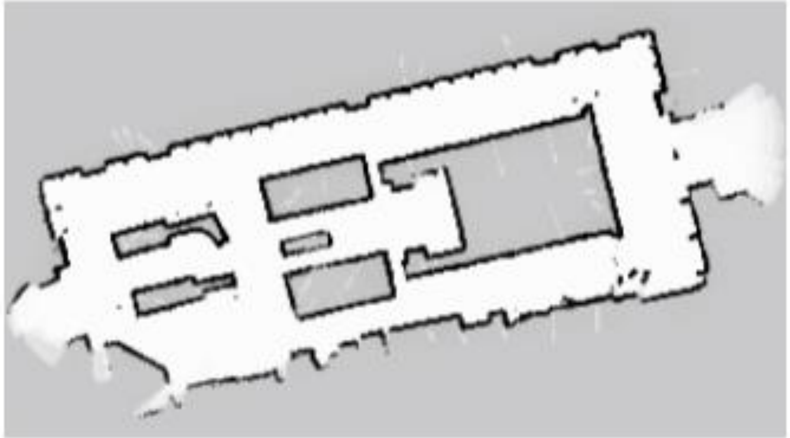
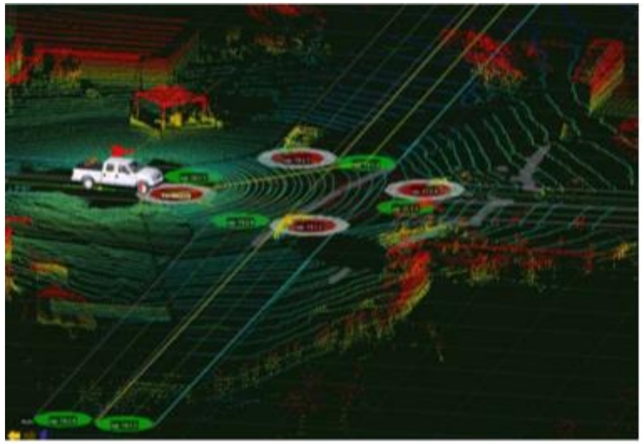
# Summary

- Random Variables are functions that map from outcomes of an experiment to the Real numbers.
- We use probability distributions (**PMF** for Discrete RVs and **PDF** for Continuous RVs.) to formally define the probability of certain outcomes.
- We use ***Joint*** probability distributions and ***Marginalization/Conditionalization*** to analyze the relationship between multiple random variables.
- Bayes rule allows us to compute probabilities that are hard to assess otherwise.

BYU Electrical & Computer Engineering





# PROBABILITY REVIEW - EXP/VAR/COVARIANCE

## ECEN 633: Robotic Localization and Mapping

Slides Based on probabilistic-robotics.org and a slide deck by Ryan Eustice.

# Common Statistics

- Expectation
- Variance
- Co-variance

# Expectation

- Weighted average according to probability

$$\mu_x = E[x] = \int_{-\infty}^{\infty} x p(x) dx$$

- Basic properties of expectation

$$E[\alpha] = \alpha$$
$$E[\alpha x] = \alpha E[x]$$
$$E[\alpha + x] = \alpha + E[x]$$
$$E[x + y] = E[x] + E[y]$$

# Variance & Covariance

- Average squared deviation from the mean.

- Variance

$$\sigma_x^2 = E[(x - E[x])^2]$$

- Covariance

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\Sigma = E[(\mathbf{x} - E[\mathbf{x}])(\mathbf{x} - E[\mathbf{x}])^\top] = \begin{bmatrix} \sigma_{xx}^2 & \sigma_{xy}^2 \\ \sigma_{yx}^2 & \sigma_{yy}^2 \end{bmatrix}$$

# Covariance Matrix Block Notation

$$\Sigma = \begin{bmatrix} \sigma_{xx}^2 & \sigma_{xy}^2 \\ \sigma_{yx}^2 & \sigma_{yy}^2 \end{bmatrix}$$

- Scalar Random Variables
  - Auto (covariance) or variance: $\sigma_{xx}^2 = E[(x - E[x])^2]$
  - Cross Covariance: $\sigma_{xy}^2 = E[(x - E[x])(y - E[y])]$

- Multi-variate Random Variables

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

  - Auto (covariance) :

$$\Sigma_{\mathbf{xx}} = E[(\mathbf{x} - E[\mathbf{x}])(\mathbf{x} - E[\mathbf{x}])^\top]$$

  - Cross Covariance:

$$\Sigma_{\mathbf{xy}} = E[(\mathbf{x} - E[\mathbf{x}])(\mathbf{y} - E[\mathbf{y}])^\top]$$

$$\Sigma = \begin{bmatrix} \Sigma_{\mathbf{xx}} & \Sigma_{\mathbf{xy}} \\ \Sigma_{\mathbf{yx}} & \Sigma_{\mathbf{yy}} \end{bmatrix}$$

# Correlation Coefficient

- The correlation coefficient is defined as:

$$\rho_{xy} = \frac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y} \qquad\qquad |\rho_{xy}| \leq 1$$

- Covariance matrix in terms of correlation coefficients

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \rho_{21}\sigma_2\sigma_1 & \sigma_2^2 & \cdots & \rho_{2n}\sigma_2\sigma_n \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1}\sigma_n\sigma_1 & \rho_{n2}\sigma_n\sigma_2 & \cdots & \sigma_n^2 \end{bmatrix}$$

# Properties of the Covariance Matrix

$$\Sigma = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

- Symmetric $B = C^\top$ why?

- Positive (semi) definite

$$\mathbf{a}^\top \Sigma \mathbf{a} \geq 0$$

# Positive (Semi) Definite

- **Def:** A matrix is positive semi definite if: $\mathbf{a}^\top \Sigma \mathbf{a} \geq 0 \quad \forall \mathbf{a}$
- A matrix is positive definite if
  - 1. it is symmetric and
  - 2. all its eigenvalues are positive.

# Eigen Decomposition

- A (non-zero) vector **v** of dimension N is an **eigenvector** of a square N x N matrix **A** if it satisfies the following equation:

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

$\lambda$ is called an **eigenvalue**

- Let **A** be a square N x N matrix with N linearly independent eigenvectors q_i. Then **A** can be factorized as:

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}$$

Where the columns of $\mathbf{Q}$ are the eigenvectors q_i and $\mathbf{\Lambda}$ is a diagonal matrix containing the eigen values of **A**.

# Properties of the Covariance Matrix

$$\Sigma = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

- Symmetric $B = C^\top$ why?
- Positive (semi) definite

$$\mathbf{a}^\top \Sigma \mathbf{a} \geq 0 \quad \text{why?}$$

- Determinant → Volume of uncertainty (Product of the Eigenvalues)
- Inverse is also positive definite

1. All Eigenvalues are Non-negative

2. If all upper left determinates are non-negative => matrix is positive (semi) definite.

Determinant:

$$|\Sigma| = a * d - b * c$$

# Joint Expectation

$$E[f(x,y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)p(x,y)dxdy$$

- Uncorrelated: $E[xy] = E[x]E[y]$
  - Independence $\rightarrow$ Uncorrelated
  - Uncorrelated $\not\rightarrow$ Independence
    - e.g.

$$p(x,y) = \tfrac{1}{4}\delta(x,y-1) + \tfrac{1}{4}\delta(x,y+1) + \tfrac{1}{4}\delta(x-1,y) + \tfrac{1}{4}\delta(x+1,y)$$

- Conditional Expectation: $E[x|y] = \int_{-\infty}^{\infty} x\, p(x|y)dx$
  - $E[x|y] = E[x]$ implies neither independence nor uncorrelatedness
    - e.g. $p(x,y) = \dfrac{1}{3}\delta(x,y+1) + \dfrac{1}{3}\delta(x+1,y) + \dfrac{1}{3}\delta(x-1,y)$

# Expectation Exercise

- We know that:

$$\Sigma_{\mathbf{xx}} = E[(\mathbf{x} - E[\mathbf{x}])(\mathbf{x} - E[\mathbf{x}])^\top]$$

- Suppose we measure a bunch of samples of $\mathbf{x}$. We compute the first and second moments of $\mathbf{x}$, i.e.,

$$M_{\mathbf{x}} = \sum \mathbf{x} \qquad M_{\mathbf{xx}} = \sum \mathbf{x}\mathbf{x}^\top$$

- How do we compute $\sum$ using only these moments and the number of samples?

$$\begin{align}
\Sigma_{\mathbf{xx}} &= E[\mathbf{x}\mathbf{x}^T - \mathbf{x}E[\mathbf{x}^\top] - E[\mathbf{x}]\mathbf{x}^\top + E[\mathbf{x}]E[\mathbf{x}^\top]\,] \tag{1}\\
&= E[\mathbf{x}\mathbf{x}^T] - E[\mathbf{x}]E[\mathbf{x}^\top] - E[\mathbf{x}]E[\mathbf{x}^\top] + E[\mathbf{x}]E[\mathbf{x}^\top] \tag{2}\\
&= E[\mathbf{x}\mathbf{x}^T] - E[\mathbf{x}]E[\mathbf{x}^\top] \tag{3}\\
&= \frac{M_{xx}}{N} - \frac{M_x M_x^\top}{N} \tag{4}
\end{align}$$

# Projecting Covariances

- Suppose I know $\mathbf{x} \sim \mu_{\mathbf{x}}, \Sigma_{\mathbf{x}}$
- How do we handle $\mathbf{y} = \mathrm{A}\mathbf{x} + \mathbf{b}$ **???**

$$\Sigma_{\mathbf{yy}} = E[(\mathbf{y} - E[\mathbf{y}])(\mathbf{y} - E[\mathbf{y}])^\top]$$

- (Algebra) ➔ $\Sigma_{\mathbf{yy}} = \mathrm{A}\Sigma_{\mathbf{xx}}\mathrm{A}^\top$

# Summary

- Common Statistics:
  - Expectation
  - Variance/Covariance
- A valid Covariance Matrix is:
  - Symmetric
  - Positive Semi-Definite
- Independence implies uncorrelated, but the opposite is **not** true.
- Expectation and Covariance can be projected easily through linear functions.