# Guiding Questions:

Utility: How good is a state or action

Probability: How likely a state or action is to occur

# Guiding Questions:

1. How do we **encode relationships** between random variables?

# Guiding Questions:

1. How do we **encode relationships** between random variables?
2. How do we **infer** something about one random variable given the value of another related one?

# Plausibility and Probability

$A, B$ $\qquad A \prec B \quad A \sim B \quad A \preceq B$

Universal Comparability : Exactly one holds

$A \prec B$
$A \succ B$
$A \sim B$

Transitivity : if $A \preceq B$ and $B \preceq C$
the $A \preceq C$

$\exists$ real-valued $P$ s.t.

$P(A) > P(B)$ iff $A \succ B$
$P(A) = P(B)$ iff $A \sim B$

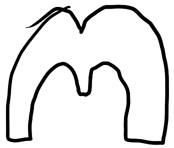
Full story: https://projecteuclid.org/journals/statistical-science/volume-1/issue-3/The-Axioms-of-Subjective-Probability/10.1214/ss/1177013611.full

**slide not on exam**

# What is a Random Variable?

today only!
Capital letters

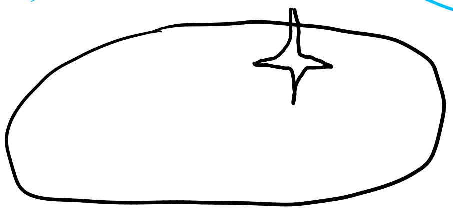
R.V. $X$



Variable

- finite set of values
- probability for each value

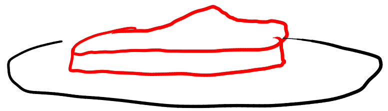
$P(X=1) = 0.5$



Variable

- continuous / discrete
- related to other R.V.s

$P(X|Y)$



$(\Omega, \Sigma, \mu)$

$X: \Omega \to E$

this section not on exam

# Vocabulary/Notation

# Vocabulary/Notation

| Term | Definition | Coinflip Example | Uniform Example |
| --- | --- | --- | --- |

# Vocabulary/Notation

| Term | Definition | Coinflip Example<br>$\mathrm{Bernoulli}(0.5)$ | Uniform Example |
|---|---|---|---|

# Vocabulary/Notation

| Term | Definition | $\mathrm{Bernoulli}(0.5)$<br>Coinflip Example | $\mathcal{U}(0,1)$<br>Uniform Example |
|---|---|---|---|

# Vocabulary/Notation

| Term | Definition | Coinflip Example<br>$\text{Bernoulli}(0.5)$ | Uniform Example<br>$\mathcal{U}(0,1)$ |
|---|---|---|---|
| support($X$) | | | |

# Vocabulary/Notation

| Term | Definition | $\text{Bernoulli}(0.5)$<br>**Coinflip Example** | $\mathcal{U}(0,1)$<br>**Uniform Example** |
|---|---|---|---|
| support($X$) | All the values that $X$ can take | | |

# Vocabulary/Notation

| Term | Definition | $\mathrm{Bernoulli}(0.5)$<br>Coinflip Example | $\mathcal{U}(0,1)$<br>Uniform Example |
|---|---|---|---|
| support($X$)<br>$x \in X$ | All the values that $X$ can take | | |

# Vocabulary/Notation

| Term | Definition | $\text{Bernoulli}(0.5)$<br>**Coinflip Example** | $\mathcal{U}(0,1)$<br>**Uniform Example** |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0,1]$ | All the values that $X$ can take | | |

# Vocabulary/Notation

| Term | Definition | $\text{Bernoulli}(0.5)$ <br> **Coinflip Example** | $\mathcal{U}(0,1)$ <br> **Uniform Example** |
|---|---|---|---|
| support($X$) <br> $x \in X$ <br> $X \in [0,1]$ | All the values that $X$ can take | $\{h, t\}$ or $\{0, 1\}$ | |

# Vocabulary/Notation

| Term | Definition | $\text{Bernoulli}(0.5)$<br>**Coinflip Example** | $\mathcal{U}(0,1)$<br>**Uniform Example** |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0,1]$ | All the values that $X$ can take | $\{h,t\}$ or $\{0,1\}$ | $[0,1]$ |

# Vocabulary/Notation

| **Term** | **Definition** | $\mathrm{Bernoulli}(0.5)$ **Coinflip Example** | $\mathcal{U}(0,1)$ **Uniform Example** |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0,1]$ | All the values that $X$ can take | $\{h,t\}$ or $\{0,1\}$ | $[0,1]$ |
| Distribution | | | |

# Vocabulary/Notation

| Term | Definition | $\text{Bernoulli}(0.5)$ Coinflip Example | $\mathcal{U}(0,1)$ Uniform Example |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0,1]$ | All the values that $X$ can take | $\{h,t\}$ or $\{0,1\}$ | $[0,1]$ |
| Distribution<br>• Discrete: PMF<br>• Continuous: PDF | | | |

# Vocabulary/Notation

| Term | Definition | $\text{Bernoulli}(0.5)$<br>**Coinflip Example** | $\mathcal{U}(0, 1)$<br>**Uniform Example** |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0, 1]$ | All the values that $X$ can take | $\{h, t\}$ or $\{0, 1\}$ | $[0, 1]$ |
| Distribution<br>• Discrete: PMF<br>• Continuous: PDF | Maps each value in the support to a real number indicating its probability | | |

# Vocabulary/Notation

| Term | Definition | Coinflip Example $\mathrm{Bernoulli}(0.5)$ | Uniform Example $\mathcal{U}(0,1)$ |
| --- | --- | --- | --- |
| support($X$)<br>$x \in X$<br>$X \in [0,1]$ | All the values that $X$ can take | $\{h, t\}$ or $\{0, 1\}$ | $[0,1]$ |
| Distribution<br>• Discrete: PMF<br>• Continuous: PDF | Maps each value in the support to a real number indicating its probability | $P(X=1)=0.5$<br>$P(X=0)=0.5$ | |

# Vocabulary/Notation

| Term | Definition | Coinflip Example ($\text{Bernoulli}(0.5)$) | Uniform Example ($\mathcal{U}(0,1)$) |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0,1]$ | All the values that $X$ can take | $\{h, t\}$ or $\{0, 1\}$ | $[0, 1]$ |
| Distribution<br>• Discrete: PMF<br>• Continuous: PDF | Maps each value in the support to a real number indicating its probability | $P(X=1) = 0.5$<br>$P(X=0) = 0.5$<br>$P(X)$ is a table | |

| X | P(X) |
|---|---|
| 0 | 0.5 |
| 1 | 0.5 |

# Vocabulary/Notation

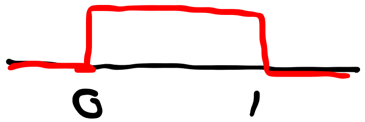
| Term | Definition | $\text{Bernoulli}(0.5)$ **Coinflip Example** | $\mathcal{U}(0, 1)$ **Uniform Example** |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0, 1]$ | All the values that $X$ can take | $\{h, t\}$ or $\{0, 1\}$ | $[0, 1]$ |
| Distribution<br>• Discrete: PMF<br>• Continuous: PDF | Maps each value in the support to a real number indicating its probability | $P(X = 1) = 0.5$<br>$P(X = 0) = 0.5$<br>$P(X)$ is a table | $p(x) = \begin{cases} 1 \text{ if } x \in [0, 1] \\ 0 \text{ o.w.} \end{cases}$ |

| **X** | **P(X)** |
|---|---|
| 0 | 0.5 |
| 1 | 0.5 |

# Vocabulary/Notation

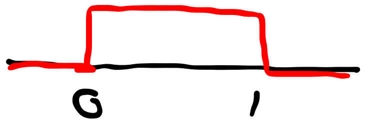
| **Term** | **Definition** | $\mathrm{Bernoulli}(0.5)$ **Coinflip Example** | $\mathcal{U}(0,1)$ **Uniform Example** |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0,1]$ | All the values that $X$ can take | $\{h, t\}$ or $\{0, 1\}$ | $[0, 1]$ |
| Distribution<br>• Discrete: PMF<br>• Continuous: PDF | Maps each value in the support to a real number indicating its probability | $P(X=1) = 0.5$<br>$P(X=0) = 0.5$<br>$P(X)$ is a table | $p(x) = \begin{cases} 1 \text{ if } x \in [0,1] \\ 0 \text{ o.w.} \end{cases}$ |

| X | P(X) |
|---|---|
| 0 | 0.5 |
| 1 | 0.5 |



# Vocabulary/Notation

| Term | Definition | $\text{Bernoulli}(0.5)$<br>Coinflip Example | $\mathcal{U}(0, 1)$<br>Uniform Example |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0, 1]$ | All the values that $X$ can take | $\{h, t\}$ or $\{0, 1\}$ | $[0, 1]$ |
| Distribution<br>• Discrete: PMF<br>• Continuous: PDF | Maps each value in the support to a real number indicating its probability | $P(X = 1) = 0.5$<br>$P(X = 0) = 0.5$<br>$P(X)$ is a table | $p(x) = \begin{cases} 1 \text{ if } x \in [0, 1] \\ 0 \text{ o.w.} \end{cases}$<br>$P(X = 1) =$ |

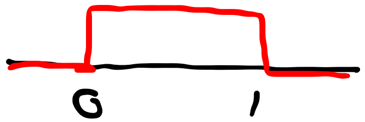



| X | P(X) |
|---|---|
| 0 | 0.5 |
| 1 | 0.5 |

# Vocabulary/Notation

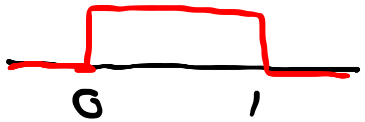
| Term | Definition | Coinflip Example | Uniform Example |
|---|---|---|---|
| | | $\text{Bernoulli}(0.5)$ | $\mathcal{U}(0,1)$ |
| support($X$)<br>$x \in X$<br>$X \in [0,1]$ | All the values that $X$ can take | $\{h, t\}$ or $\{0,1\}$ | $[0,1]$ |
| Distribution<br>• Discrete: PMF<br>• Continuous: PDF | Maps each value in the support to a real number indicating its probability | $P(X=1) = 0.5$<br>$P(X=0) = 0.5$<br>$P(X)$ is a table | $p(x) = \begin{cases} 1 \text{ if } x \in [0,1] \\ 0 \text{ o.w.} \end{cases}$<br>$P(X=1) = 0$ |

| X | P(X) |
|---|---|
| 0 | 0.5 |
| 1 | 0.5 |



# Vocabulary/Notation

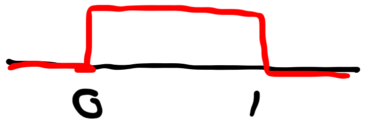
| Term | Definition | $\text{Bernoulli}(0.5)$ Coinflip Example | $\mathcal{U}(0,1)$ Uniform Example |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0,1]$ | All the values that $X$ can take | $\{h, t\}$ or $\{0, 1\}$ | $[0,1]$ |
| Distribution<br>• Discrete: PMF<br>• Continuous: PDF | Maps each value in the support to a real number indicating its probability | $P(X=1)=0.5$<br>$P(X=0)=0.5$<br>$P(X)$ is a table | $p(x) = \begin{cases} 1 \text{ if } x \in [0,1] \\ 0 \text{ o.w.} \end{cases}$<br>$P(X=1) = 0$<br>$P(X \in [a,b]) = \int_a^b p(x)dx$ |




| X | P(X) |
|---|---|
| 0 | 0.5 |
| 1 | 0.5 |

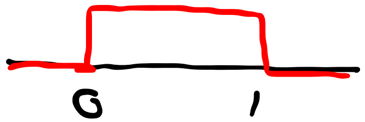
# Vocabulary/Notation

| Term | Definition | $\text{Bernoulli}(0.5)$<br>**Coinflip Example** | $\mathcal{U}(0,1)$<br>**Uniform Example** |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0,1]$ | All the values that $X$ can take | $\{h,t\}$ or $\{0,1\}$ | $[0,1]$ |
| Distribution<br>• Discrete: PMF<br>• Continuous: PDF | Maps each value in the support to a real number indicating its probability | $P(X=1)=0.5$<br>$P(X=0)=0.5$<br>$P(X)$ is a table | $p(x) = \begin{cases} 1 \text{ if } x \in [0,1] \\ 0 \text{ o.w.} \end{cases}$<br>$P(X=1) = 0$<br>$P(X \in [a,b]) = \int_a^b p(x)dx$ |
| Expectation | | | |




| X | P(X) |
|---|---|
| 0 | 0.5 |
| 1 | 0.5 |

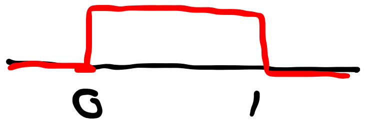
# Vocabulary/Notation

<table>
<tr><th>Term</th><th>Definition</th><th>$\text{Bernoulli}(0.5)$<br>Coinflip Example</th><th>$\mathcal{U}(0,1)$<br>Uniform Example</th></tr>
<tr><td>support($X$)<br>$x \in X$<br>$X \in [0,1]$</td><td>All the values that $X$ can take</td><td>$\{h,t\}$ or $\{0,1\}$</td><td>$[0,1]$<br><br></td></tr>
<tr><td>Distribution<br>• Discrete: PMF<br>• Continuous: PDF</td><td>Maps each value in the support to a real number indicating its probability</td><td>$P(X=1)=0.5$<br>$P(X=0)=0.5$<br>$P(X)$ is a table<br><table><tr><th>X</th><th>P(X)</th></tr><tr><td>0</td><td>0.5</td></tr><tr><td>1</td><td>0.5</td></tr></table></td><td>$p(x) = \begin{cases} 1 \text{ if } x \in [0,1] \\ 0 \text{ o.w.} \end{cases}$<br>$P(X=1) = 0$<br>$P(X \in [a,b]) = \int_a^b p(x)dx$</td></tr>
<tr><td>Expectation</td><td>Single representative value of the random variable, "mean"</td><td></td><td></td></tr>
</table>

# Vocabulary/Notation

| Term | Definition | Coinflip Example $\text{Bernoulli}(0.5)$ | Uniform Example $\mathcal{U}(0,1)$ |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0,1]$ | All the values that $X$ can take | $\{h,t\}$ or $\{0,1\}$ | $[0,1]$ |
| Distribution<br>• Discrete: PMF<br>• Continuous: PDF | Maps each value in the support to a real number indicating its probability | $P(X=1)=0.5$<br>$P(X=0)=0.5$<br>$P(X)$ is a table | $p(x) = \begin{cases} 1 \text{ if } x \in [0,1] \\ 0 \text{ o.w.} \end{cases}$<br>$P(X=1) = 0$<br>$P(X \in [a,b]) = \int_a^b p(x)dx$ |
| Expectation<br>$E[X]$ | Single representative value of the random variable, "mean" | | |

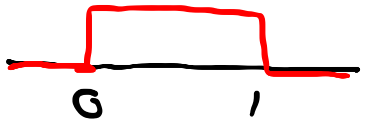
| X | P(X) |
|---|---|
| 0 | 0.5 |
| 1 | 0.5 |



# Vocabulary/Notation

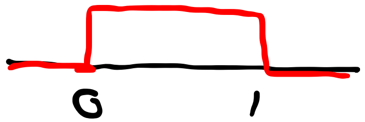
| **Term** | **Definition** | $\text{Bernoulli}(0.5)$<br>**Coinflip Example** | $\mathcal{U}(0,1)$<br>**Uniform Example** |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0,1]$ | All the values that $X$ can take | $\{h,t\}$ or $\{0,1\}$ | $[0,1]$ |
| Distribution<br>• Discrete: PMF<br>• Continuous: PDF | Maps each value in the support to a real number indicating its probability | $P(X=1)=0.5$<br>$P(X=0)=0.5$<br>$P(X)$ is a table | $p(x) = \begin{cases} 1 \text{ if } x \in [0,1] \\ 0 \text{ o.w.} \end{cases}$<br>$P(X=1)=0$<br>$P(X \in [a,b]) = \int_a^b p(x)dx$ |
| Expectation<br>$E[X]$ | Single representative value of the random variable, "mean" | $E[X] = \sum_{x \in X} xP(x)$ | |

| X | P(X) |
|---|---|
| 0 | 0.5 |
| 1 | 0.5 |



# Vocabulary/Notation

| Term | Definition | Coinflip Example ($\text{Bernoulli}(0.5)$) | Uniform Example ($\mathcal{U}(0,1)$) |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0,1]$ | All the values that $X$ can take | $\{h, t\}$ or $\{0, 1\}$ | $[0, 1]$ |




**Distribution**

- Discrete: PMF
- Continuous: PDF

Maps each value in the support to a real number indicating its probability

Coinflip Example:

$$P(X=1) = 0.5$$
$$P(X=0) = 0.5$$

$P(X)$ is a table

| X | P(X) |
|---|---|
| 0 | 0.5 |
| 1 | 0.5 |

Uniform Example:

$$p(x) = \begin{cases} 1 \text{ if } x \in [0,1] \\ 0 \text{ o.w.} \end{cases}$$

$$P(X=1) = 0$$

$$P(X \in [a,b]) = \int_a^b p(x)dx$$

**Expectation**

$E[X]$

Single representative value of the random variable, "mean"

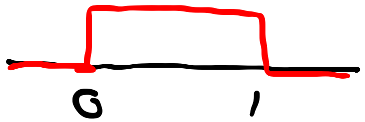
Coinflip Example:

$$E[X] = \sum_{x \in X} xP(x)$$
$$= 0.5$$

# Vocabulary/Notation

| Term | Definition | $\text{Bernoulli}(0.5)$ **Coinflip Example** | $\mathcal{U}(0,1)$ **Uniform Example** |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0,1]$ | All the values that $X$ can take | $\{h,t\}$ or $\{0,1\}$ | $[0,1]$ |
| Distribution<br>• Discrete: PMF<br>• Continuous: PDF | Maps each value in the support to a real number indicating its probability | $P(X=1)=0.5$<br>$P(X=0)=0.5$<br>$P(X)$ is a table (see below) | $p(x) = \begin{cases} 1 \text{ if } x \in [0,1] \\ 0 \text{ o.w.} \end{cases}$<br>$P(X=1)=0$<br>$P(X \in [a,b]) = \int_a^b p(x)dx$ |
| Expectation<br>$E[X]$ | Single representative value of the random variable, "mean" | $E[X] = \sum_{x \in X} xP(x)$<br>$= 0.5$ | $E[X] = \int_{x \in X} xp(x)dx$ |

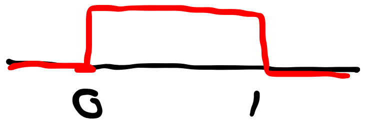
| X | P(X) |
|---|---|
| 0 | 0.5 |
| 1 | 0.5 |



# Vocabulary/Notation

| Term | Definition | Coinflip Example ($\text{Bernoulli}(0.5)$) | Uniform Example ($\mathcal{U}(0,1)$) |
|---|---|---|---|
| support($X$)<br>$x \in X$<br>$X \in [0,1]$ | All the values that $X$ can take | $\{h,t\}$ or $\{0,1\}$ | $[0,1]$ |
| Distribution<br>• Discrete: PMF<br>• Continuous: PDF | Maps each value in the support to a real number indicating its probability | $P(X=1)=0.5$<br>$P(X=0)=0.5$<br>$P(X)$ is a table (see below) | $p(x) = \begin{cases} 1 \text{ if } x \in [0,1] \\ 0 \text{ o.w.} \end{cases}$<br>$P(X=1) = 0$<br>$P(X \in [a,b]) = \int_a^b p(x)dx$ |
| Expectation<br>$E[X]$ | Single representative value of the random variable, "mean" | $E[X] = \sum_{x \in X} xP(x)$<br>$= 0.5$ | $E[X] = \int_{x \in X} xp(x)dx$<br>$= 0.5$ |

| X | P(X) |
|---|---|
| 0 | 0.5 |
| 1 | 0.5 |



# Distributions of related R.V.s

# Distributions of related R.V.s

**Joint Distribution**

# Distributions of related R.V.s

**Joint Distribution**

$$P(X, Y, Z)$$

# Distributions of related R.V.s

## Joint Distribution

$$P(X, Y, Z)$$

| $X$ | $Y$ | $Z$ | $P(X,Y,Z)$ |
|---|---|---|---|
| 0 | 0 | 0 | 0.08 |
| 0 | 0 | 1 | 0.31 |
| 0 | 1 | 0 | 0.09 |
| 0 | 1 | 1 | 0.37 |
| 1 | 0 | 0 | 0.01 |
| 1 | 0 | 1 | 0.05 |
| 1 | 1 | 0 | 0.02 |
| 1 | 1 | 1 | 0.07 |

# Distributions of related R.V.s

## Joint Distribution

$$P(X, Y, Z)$$

| $X$ | $Y$ | $Z$ | $P(X,Y,Z)$ |
|---|---|---|---|
| 0 | 0 | 0 | 0.08 |
| 0 | 0 | 1 | 0.31 |
| 0 | 1 | 0 | 0.09 |
| 0 | 1 | 1 | 0.37 |
| 1 | 0 | 0 | 0.01 |
| 1 | 0 | 1 | 0.05 |
| 1 | 1 | 0 | 0.02 |
| 1 | 1 | 1 | 0.07 |

## Conditional Distribution

# Distributions of related R.V.s

## Joint Distribution

$$P(X, Y, Z)$$

| $X$ | $Y$ | $Z$ | $P(X,Y,Z)$ |
|---|---|---|---|
| 0 | 0 | 0 | 0.08 |
| 0 | 0 | 1 | 0.31 |
| 0 | 1 | 0 | 0.09 |
| 0 | 1 | 1 | 0.37 |
| 1 | 0 | 0 | 0.01 |
| 1 | 0 | 1 | 0.05 |
| 1 | 1 | 0 | 0.02 |
| 1 | 1 | 1 | 0.07 |

## Conditional Distribution

$$P(X \mid Y, Z)$$

# Distributions of related R.V.s

## Joint Distribution

$$P(X, Y, Z)$$

| $X$ | $Y$ | $Z$ | $P(X, Y, Z)$ |
|---|---|---|---|
| 0 | 0 | 0 | 0.08 |
| 0 | 0 | 1 | 0.31 |
| 0 | 1 | 0 | 0.09 |
| 0 | 1 | 1 | 0.37 |
| 1 | 0 | 0 | 0.01 |
| 1 | 0 | 1 | 0.05 |
| 1 | 1 | 0 | 0.02 |
| 1 | 1 | 1 | 0.07 |

## Conditional Distribution

$$P(X \mid Y, Z)$$

(Distribution - valued function)

| $X$ | $P(X \mid Y{=}1, Z{=}1)$ |
|---|---|
| 0 | 0.84 |
| 1 | 0.16 |

# Distributions of related R.V.s

## Joint Distribution

$$P(X, Y, Z)$$

| $X$ | $Y$ | $Z$ | $P(X,Y,Z)$ |
|---|---|---|---|
| 0 | 0 | 0 | 0.08 |
| 0 | 0 | 1 | 0.31 |
| 0 | 1 | 0 | 0.09 |
| 0 | 1 | 1 | 0.37 |
| 1 | 0 | 0 | 0.01 |
| 1 | 0 | 1 | 0.05 |
| 1 | 1 | 0 | 0.02 |
| 1 | 1 | 1 | 0.07 |

## Conditional Distribution

$$P(X \mid Y, Z)$$

(Distribution - valued function)

| $X$ | $P(X \mid Y=1, Z=1)$ |
|---|---|
| 0 | 0.84 |
| 1 | 0.16 |

## Marginal Distribution

# Distributions of related R.V.s

## Joint Distribution

$$P(X, Y, Z)$$

| $X$ | $Y$ | $Z$ | $P(X,Y,Z)$ |
|---|---|---|---|
| 0 | 0 | 0 | 0.08 |
| 0 | 0 | 1 | 0.31 |
| 0 | 1 | 0 | 0.09 |
| 0 | 1 | 1 | 0.37 |
| 1 | 0 | 0 | 0.01 |
| 1 | 0 | 1 | 0.05 |
| 1 | 1 | 0 | 0.02 |
| 1 | 1 | 1 | 0.07 |

## Conditional Distribution

$$P(X \mid Y, Z)$$

(Distribution - valued function)

| $X$ | $P(X \mid Y=1, Z=1)$ |
|---|---|
| 0 | 0.84 |
| 1 | 0.16 |

## Marginal Distribution

$$P(X)\ P(Y)\ P(Z)$$

# Distributions of related R.V.s

## Joint Distribution

$$P(X, Y, Z)$$

| $X$ | $Y$ | $Z$ | $P(X,Y,Z)$ |
|---|---|---|---|
| 0 | 0 | 0 | 0.08 |
| 0 | 0 | 1 | 0.31 |
| 0 | 1 | 0 | 0.09 |
| 0 | 1 | 1 | 0.37 |
| 1 | 0 | 0 | 0.01 |
| 1 | 0 | 1 | 0.05 |
| 1 | 1 | 0 | 0.02 |
| 1 | 1 | 1 | 0.07 |

## Conditional Distribution

$$P(X \mid Y, Z)$$

(Distribution - valued function)

| $X$ | $P(X \mid Y=1, Z=1)$ |
|---|---|
| 0 | 0.84 |
| 1 | 0.16 |

## Marginal Distribution

$$P(X)\ P(Y)\ P(Z)$$

| $X$ | $P(X)$ |
|---|---|
| 0 | 0.85 |
| 1 | 0.15 |

| $Y$ | $P(Y)$ |
|---|---|
| 0 | 0.45 |
| 1 | 0.55 |

| $Z$ | $P(Z)$ |
|---|---|
| 0 | 0.20 |
| 1 | 0.80 |

# Distributions of related R.V.s

**Joint Distribution**

$$P(X, Y, Z)$$

**Conditional Distribution**

$$P(X \mid Y, Z)$$

**Marginal Distribution**

$$P(X)\ P(Y)\ P(Z)$$

# Distributions of related R.V.s

**Joint Distribution**

$$P(X, Y, Z)$$

**Conditional Distribution**

$$P(X \mid Y, Z)$$

**Marginal Distribution**

$$P(X)\ P(Y)\ P(Z)$$

3 Rules

# Distributions of related R.V.s

**Joint Distribution**

$P(X, Y, Z)$

**Conditional Distribution**

$P(X \mid Y, Z)$

**Marginal Distribution**

$P(X)\ P(Y)\ P(Z)$

3 Rules

(Burrito-level)

# Distributions of related R.V.s

**Joint Distribution**

$P(X, Y, Z)$

**Conditional Distribution**

$P(X \mid Y, Z)$

**Marginal Distribution**

$P(X)\ P(Y)\ P(Z)$

3 Rules

(Burrito-level)

(Filet Minion Level: Axioms of Probability)

**AXIOM 1. STRUCTURE OF UNKNOWN REAL NUMBERS AND PLAUSIBLE VALUE.** *We assume a set $T$ of unknown numbers is a partially ordered commutative algebra over $\mathbb{R}$ with identity, 1.*

*We assume in addition a given sub-Boolean algebra $E$ of $E(T)$ with $0, 1 \in E$ and denote by $E_0$ the set of non-zero members of $E$. We assume that the partial ordering in $E(T)$ as a Boolean algebra coincides with the ordering that $E(T)$ inherits from the algebra $T$. Finally, we assume a function $PV : T \times E_0 \to \mathbb{R}$, called* **PLAUSIBLE VALUE**, *whose value on the pair $(x, e)$ is denoted $PV(x|e)$.*

**AXIOM 2. STRONG RESCALING FOR PLAUSIBLE VALUE.** *If $a, b$ belong to $\mathbb{R}$, if $x$ belongs to $T$, and if $e$ belongs to $E_0$, then*

$$PV(ax + b|e) = aPV(x|e) + b. \tag{2}$$

**AXIOM 3. ORDER CONSISTENCY FOR PLAUSIBLE VALUE.** *If $x, y \in T$ and if $e \in E_0$, implies that $x \leq y$, then $PV(x|e) \leq PV(y|e)$.*

Notice that if $e \in E(T)$, then $0 \leq e \leq 1$, in $T$, as it is true in the lattice ordering of $E(T)$.

**AXIOM 4. THE COX AXIOM FOR PLAUSIBLE VALUE:** *If $e, c$ are fixed in $E$, with $ec \in E_0$, if $x_1, x_2$ are in $T$, if $PV(x_1|ec) = PV(x_2|ec)$, then $PV(x_1e|c) = PV(x_2e|c)$. That is, we assume that as a function of $x$, the plausible value $PV(xe|c)$ depends only on $PV(x|ec)$.*

**AXIOM 5. RESTRICTED ADDITIVITY OF PLAUSIBLE VALUE.** *For each fixed $y \in T$ and $e \in E_0$, the plausible value $PV(x + y|e)$ as a function of $x \in T$ depends only on $PV(x|e)$, which is to say that if $x_1, x_2 \in T$ and $PV(x_1|e) = PV(x_2|e)$, then $PV(x_1 + y|e) = PV(x_2 + y|e)$.*

not on exam

# Distributions of related R.V.s

**Joint Distribution**

$P(X, Y, Z)$

**Conditional Distribution**

$P(X \mid Y, Z)$

**Marginal Distribution**

$P(X)\ P(Y)\ P(Z)$

## 3 Rules

(Burrito-level)

1)

# Distributions of related R.V.s

**Joint Distribution**

$P(X, Y, Z)$

**Conditional Distribution**

$P(X \mid Y, Z)$

**Marginal Distribution**

$P(X)\ P(Y)\ P(Z)$

## 3 Rules

(Burrito-level)

1) a) $0 \leq P(X \mid Y) \leq 1$

# Distributions of related R.V.s

**Joint Distribution**

$P(X, Y, Z)$

**Conditional Distribution**

$P(X \mid Y, Z)$

**Marginal Distribution**

$P(X)\ P(Y)\ P(Z)$

## 3 Rules

(Burrito-level)

1) a) $0 \leq P(X \mid Y) \leq 1$

   b) $\sum_{x \in X} P(x \mid Y) = 1$

# Distributions of related R.V.s

**Joint Distribution**

$$P(X, Y, Z)$$

**Conditional Distribution**

$$P(X \mid Y, Z)$$

**Marginal Distribution**

$$P(X)\ P(Y)\ P(Z)$$

## 3 Rules

(Burrito-level)

1) a) $0 \leq P(X \mid Y) \leq 1$

   b) $\sum_{x \in X} P(x \mid Y) = 1$

2) "Law of total probability"

$$P(X) = \sum_{y \in Y} P(X, y)$$

# Distributions of related R.V.s

**Joint Distribution**

$P(X, Y, Z)$

**Conditional Distribution**

$P(X \mid Y, Z)$

**Marginal Distribution**

$P(X)\ P(Y)\ P(Z)$

## 3 Rules

(Burrito-level)

1) a) $0 \leq P(X \mid Y) \leq 1$

   b) $\sum_{x \in X} P(x \mid Y) = 1$

2) "Law of total probability"

$$P(X) = \sum_{y \in Y} P(X, y)$$

Joint $\rightarrow$ Marginal

# Distributions of related R.V.s

**Joint Distribution**

$P(X, Y, Z)$

**Conditional Distribution**

$P(X \mid Y, Z)$

**Marginal Distribution**

$P(X)\ P(Y)\ P(Z)$

## 3 Rules

(Burrito-level)

1) a) $0 \leq P(X \mid Y) \leq 1$

   b) $\sum_{x \in X} P(x \mid Y) = 1$

2) "Law of total probability"

$$P(X) = \sum_{y \in Y} P(X, y)$$

Joint $\rightarrow$ Marginal

3) Definition of Conditional Probability

$$P(X \mid Y) = \frac{P(X, Y)}{P(Y)}$$

# Distributions of related R.V.s

**Joint Distribution**

$P(X, Y, Z)$

**Conditional Distribution**

$P(X \mid Y, Z)$

**Marginal Distribution**

$P(X)\ P(Y)\ P(Z)$

## 3 Rules

(Burrito-level)

1) a) $0 \leq P(X \mid Y) \leq 1$

   b) $\sum_{x \in X} P(x \mid Y) = 1$

2) "Law of total probability"

$$P(X) = \sum_{y \in Y} P(X, y)$$

3) Definition of Conditional Probability

$$P(X \mid Y) = \frac{P(X, Y)}{P(Y)}$$

Joint $\rightarrow$ Marginal

Joint + Marginal $\rightarrow$ Conditional

# Distributions of related R.V.s

**Joint Distribution**

$P(X, Y, Z)$

**Conditional Distribution**

$P(X \mid Y, Z)$

**Marginal Distribution**

$P(X)\ P(Y)\ P(Z)$

## 3 Rules

(Burrito-level)

1) a) $0 \leq P(X \mid Y) \leq 1$

   b) $\sum_{x \in X} P(x \mid Y) = 1$

2) "Law of total probability"

$$P(X) = \sum_{y \in Y} P(X, y)$$

3) Definition of Conditional Probability

$$P(X \mid Y) = \frac{P(X, Y)}{P(Y)}$$

Joint $\rightarrow$ Marginal

Joint + Marginal $\rightarrow$ Conditional

Marginal + Conditional $\rightarrow$ Joint

$$P(X, Y) = P(X|Y)\, P(Y)$$

# Distributions of related R.V.s

**Joint Distribution** | **Conditional Distribution** | **Marginal Distribution**

$P(X, Y, Z)$

$P(X \mid Y, Z)$

$P(X)\, P(Y)\, P(Z)$

| X | Y | Z | P(X,Y,Z) |
|---|---|---|---|
| 0 | 0 | 0 | 0.08 |
| 0 | 0 | 1 | 0.31 |
| 0 | 1 | 0 | 0.09 |
| 0 | 1 | 1 | 0.37 |
| 1 | 0 | 0 | 0.01 |
| 1 | 0 | 1 | 0.05 |
| 1 | 1 | 0 | 0.02 |
| 1 | 1 | 1 | 0.07 |

$P(X) = \sum_{y \in Y, z \in Z} P(X, y, z)$

| X | P(X) |
|---|---|
| 0 | 0.08 + 0.31 + 0.09 + 0.37 = 0.85 |
| 1 | 0.15 |

$\sum_{x \in X} P(x) = 1 = 0.85 + P(X=1)$

## 3 Rules

1) a) $0 \leq P(X \mid Y) \leq 1$

   b) $\sum_{x \in X} P(x \mid Y) = 1$

2) "Law of total probability"

$$P(X) = \sum_{y \in Y} P(X, y)$$

3) Definition of Conditional Probability

$$P(X \mid Y) = \frac{P(X, Y)}{P(Y)}$$

Joint $\rightarrow$ Marginal

$P(Y=1, Z=1 \mid X=1) = \frac{P(X=1, Y=1, Z=1)}{P(X=1)} = \frac{0.07}{0.15} = 0.47$

Joint + Marginal $\rightarrow$ Conditional

Marginal + Conditional $\rightarrow$ Joint

$$P(X, Y) = P(X|Y)\, P(Y)$$

1) a) $0 \le P(X \mid Y) \le 1$

b) $\sum_{x \in X} P(x \mid Y) = 1$

2) $P(X) = \sum_{y \in Y} P(X, y)$

3) $P(X \mid Y) = \frac{P(X,Y)}{P(Y)}$

$$P(X,Y) = P(X|Y)\,P(Y)$$

# Break

$P(C=1) = P(C=1, P=1) + \underline{P(C=1, P=0)}$

$P(P=0, C=1) = P(C=1) - P(C=1, P=1)$

$0.7 \leftarrow 0.8 \qquad 0.1$

| P | C | |
|---|---|---|
| 0 | 0 | 0.1 |
| 0 | 1 | 0.7 |
| 1 | 0 | 0.1 |
| 1 | 1 | 0.1 |

- $P \in \{0, 1\}$: Powder Day
- $C \in \{0, 1\}$: Pass Clear
- 1 in 5 days is a powder day $P(P=1) = 0.2$
- The pass is clear 8 in 10 days $P(C=1) = 0.8$
- If it is a powder day, there is a 50% chance the pass is blocked $P(C=0 \mid P=1) = 0.5$

$P(C=1 \mid P=1) = 0.5$

- What is the probability that there is a powder day and the pass is clear? $P(P=1, C=1)$ 0.1
- What is the probability that the pass is blocked on a non-powder day $P(C=0 \mid P=0) = \frac{P(C=0, P=0)}{P(P=0)} = \frac{0.1}{0.8} = 0.125$

# Bayes Rule

- Know: $P(B \mid A)$, $P(A)$, $P(B)$
- Want: $P(A \mid B)$

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

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

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

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

$P(X,Y) = P(X)P(Y)$

# Independence

# Independence

Definition: $X$ and $Y$ are *independent* iff $P(X, Y) = P(X)\, P(Y)$

# Independence

Definition: $X$ and $Y$ are *independent* iff $P(X, Y) = P(X)\, P(Y)$

$X \perp Y$

# Independence

Definition: $X$ and $Y$ are *independent* iff $P(X, Y) = P(X)\, P(Y)$

$X \perp Y$

$P(X|Y) = P(X)$

# Independence

Definition: $X$ and $Y$ are *independent* iff $P(X, Y) = P(X)\, P(Y)$

$X \perp Y$

$P(X|Y) = P(X)$

Definition: $X$ and $Y$ are *conditionally independent* given $Z$ iff
$P(X, Y \mid Z) = P(X \mid Z)\, P(Y \mid Z)$

# Independence

Definition: $X$ and $Y$ are *independent* iff $P(X, Y) = P(X)\,P(Y)$

$X \perp Y$

$P(X \mid Y) = P(X)$

Definition: $X$ and $Y$ are *conditionally independent* given $Z$ iff
$P(X, Y \mid Z) = P(X \mid Z)\,P(Y \mid Z)$

$X \perp Y \mid Z$

$X \perp Y \mid Z \overset{?}{\Rightarrow} X \perp Y$ No!

$X \perp Y \overset{?}{\Rightarrow} X \perp Y \mid Z$ No!

# Rules for Continuous RVs

**Discrete**

1) a) $0 \leq P(X \mid Y) \leq 1$

   b) $\sum_{x \in X} P(x \mid Y) = 1$

2) $P(X) = \sum_{y \in Y} P(X, y)$

3) $P(X \mid Y) = \frac{P(X,Y)}{P(Y)}$

$P(X, Y) = P(X \mid Y)\, P(Y)$

**Continuous**

1)

# Rules for Continuous RVs

**Discrete**

1) a) $0 \leq P(X \mid Y) \leq 1$

   b) $\sum_{x \in X} P(x \mid Y) = 1$

2) $P(X) = \sum_{y \in Y} P(X, y)$

3) $P(X \mid Y) = \frac{P(X,Y)}{P(Y)}$

   $P(X, Y) = P(X \mid Y)\, P(Y)$

**Continuous**

1) $0 \leq p(X \mid Y)$

# Rules for Continuous RVs

**Discrete**

1) a) $0 \leq P(X \mid Y) \leq 1$

   b) $\sum_{x \in X} P(x \mid Y) = 1$

2) $P(X) = \sum_{y \in Y} P(X, y)$

3) $P(X \mid Y) = \frac{P(X,Y)}{P(Y)}$

   $P(X, Y) = P(X \mid Y)\, P(Y)$

**Continuous**

1) $0 \leq p(X \mid Y)$

   $\int_X p(x|Y)\, dx = 1$

# Rules for Continuous RVs

**Discrete**

1) a) $0 \leq P(X \mid Y) \leq 1$

   b) $\sum_{x \in X} P(x \mid Y) = 1$

2) $P(X) = \sum_{y \in Y} P(X, y)$

3) $P(X \mid Y) = \frac{P(X,Y)}{P(Y)}$

$$P(X, Y) = P(X \mid Y)\, P(Y)$$

**Continuous**

1) $0 \leq p(X \mid Y)$

$$\int_X p(x|Y)\, dx = 1$$

2)

$$p(X) = \int_Y p(X, y) dy$$

# Rules for Continuous RVs

**Discrete**

1) a) $0 \leq P(X \mid Y) \leq 1$

b) $\sum_{x \in X} P(x \mid Y) = 1$

2) $P(X) = \sum_{y \in Y} P(X, y)$

3) $P(X \mid Y) = \frac{P(X,Y)}{P(Y)}$

$$P(X, Y) = P(X \mid Y)\, P(Y)$$

**Continuous**

1) $0 \leq p(X \mid Y)$

$$\int_X p(x|Y)\, dx = 1$$

2)

$$p(X) = \int_Y p(X, y) dy$$

3) $p(X \mid Y) = \frac{p(X,Y)}{p(Y)}$

$$p(X, Y) = p(X \mid Y)\, p(Y)$$

# Multivariate Gaussian Distribution

$\mathcal{N}(\mu, \Sigma)$

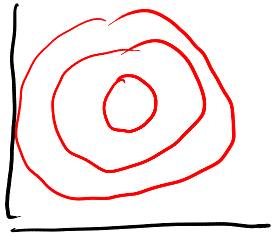
$x = [x_1, x_2]$

## Joint Distribution

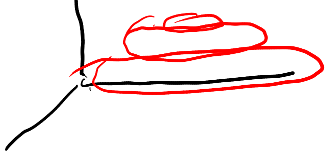
$$p(x) = \mathcal{N}(x|\mu,\Sigma)$$
$$= \frac{\exp\left(\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)}{(2\pi)^{n/2}|\Sigma|^{1/2}}$$

$\mu = [3,3]$ $\Sigma = \begin{bmatrix}3 & 0\\ 0 & 3\end{bmatrix}$ $\Sigma = \begin{bmatrix}4 & 2\\ 2 & 4\end{bmatrix}$

$x_2$



$x_1$

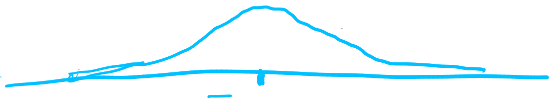


## Conditional Distribution

$$p(x_1|x_2) = \mathcal{N}(x_1|\bar{\mu},\bar{\Sigma})$$

$$\bar{\mu} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2)$$
$$\bar{\Sigma} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

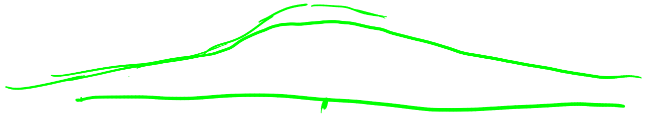
$p(x_1|x_2=4)$



$\bar{\mu} = 3.5$

## Marginal Distribution

$$p(x_1) = \mathcal{N}(x_1|\mu_1,\Sigma_{11})$$



$\mu_1 = 3$

# Guiding Questions:

# Guiding Questions:

1. How do we **encode relationships** between random variables?

# Guiding Questions:

1. How do we **encode relationships** between random variables?
2. How do we **infer** something about one random variable given the value of another related one?

Joint
Conditional
Marginal

Bayes Rule