Probability and 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

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

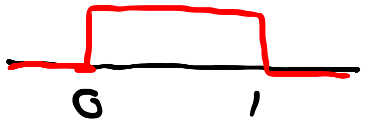
slide not on exam

What is a Random Variable?

R.V. X

this section not on exam

Vocabulary/Notation

Term	**Definition**	$\text{Bernoulli}(0.5)$ **Coinflip Example**	$\mathcal{U}(0,1)$ **Uniform Example**
support(X) $x \in X$ $X \in [0,1]$	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	$P(X=1)=0.5$ $P(X=0)=0.5$ $P(X)$ is a table <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>	$p(x) = \begin{cases} 1 \text{ if } x \in [0,1] \\ 0 \text{ o.w.} \end{cases}$ $P(X=1) = 0$ $P(X \in [a,b]) = \int_a^b p(x)dx$
Expectation $E[X]$	Single representative value of the random variable, "mean"	$E[X] = \sum_{x \in X} xP(x)$ $= 0.5$	$E[X] = \int_{x \in X} xp(x)dx$ $= 0.5$

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

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

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

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

Conditional Distribution

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

Marginal Distribution

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

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

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

Joint + Marginal $\rightarrow$ Conditional

Marginal + Conditional $\rightarrow$ Joint

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

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

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

Break

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

- What is the probability that there is a powder day and the pass is clear?
- What is the probability that the pass is blocked on a non-powder day

Bayes Rule

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

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

$X \perp Y \mid Z$

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

Joint Distribution

Conditional Distribution

Marginal Distribution

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?