today: probability review

why? → principled framework to model uncertainty

sources of uncertainty

1) intrinsic uncertainty → quantum mechanics
2) partial information/observation:  • card games  
   • rolling a dice → don't know the initial conditions
3) incomplete modeling of a complex phenomenon  
   example: "most birds can fly"  
   → simple rule is advantage but then yields uncertainty

notation: \( X_1, X_2, X_3 \) \( X, Y, Z \)

→ random variables
   (usually real-valued)
\( X_1, X_2, X_3 \) \( x, y, z \)

→ their "realizations"
$X$ a random variable → represents an uncertain quantity

$X = x$ represents the "event" that $X$ takes the value $x$

$\Omega$ → the sample space of "elementary events"

Possible values for my R.V.

e.g. $X$ is result of a die throw

$\Omega = \{1, 2, 3, 4, 5, 6\}$

two types of R.V.: discrete where $\Omega$ is countable

continous ($\Omega$ is uncountable)

[assume $\Omega$ is countable]

R.V. $X$ is characterized by a probability mass function (pmf)

$p(x)$ for $x \in \Omega$

st. $\int p(x) dx = 0 \forall x$

$\sum_{x \in \Omega} p(x) = 1$

a probability distribution $P$ is a mapping $P: \Omega \rightarrow [0, 1]$
is a mapping \( P : \mathcal{P}(\Omega) \rightarrow [0,1] \)

\( \mathcal{P}(\Omega) = \mathcal{E} \) set of all subsets of \( \Omega \)

- set of "events"

('\( \mathcal{E} \)-field' in measure theory needed when \( \Omega \) uncountable)

**Kolmogorov axioms**

\[
\begin{align*}
P(E) &> 0 \quad \forall E \in \mathcal{E} \\
P(\Omega) & = 1 \\
P(\bigcup_{i=1}^{\infty} E_i) & = \sum_{i=1}^{\infty} P(E_i) \quad \text{when } E_i \text{ are disjoint}
\end{align*}
\]

for discrete R.V. \( P(E) = \sum_{x \in E} p(x) \)

continuous R.V. is characterized by a probability density function (pdf) \( p(x) \)

\[
\begin{align*}
p(x) &> 0 \quad \forall x \in \Omega \\
p \text{ is integrable and } & \int_{\Omega} p(x) \, dx = 1
\end{align*}
\]

\( \Omega = \mathbb{R} : \quad P([a,b]) = \int_{a}^{b} p(x) \, dx \)

"Has measure 1"
Sidenote: can change on countable # of pts without changing "anything".

notion: $p_X(x)$, $p_Y(y)$
$p_Y(z)$ when $z = y$

Recap:

discrete R.V. $X$; pmf $p_X(x)$
$P_x X = x^2 = p(x)$

cts R.V. $X$; pdf $p(x)$
"$P_x X \in x \pm dx \stackrel{?}{=} p(x)dx$"

$P_x X \in x \pm \frac{\sigma}{\sqrt{2}} = \int_{x-\frac{\sigma}{\sqrt{2}}}^{x+\frac{\sigma}{\sqrt{2}}} p(u)du$

joint, marginal, etc. (multivariate R.V.)
$Z = (X,Y)$, $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$

pmf of $Z$ is "joint pmf" on $X \times Y$
$p(x, y) = P_x X = x, Y = y$
$$p(x, y) = P\{X = x, Y = y\}$$

Joint distribution

can represent elementary events as table

<table>
<thead>
<tr>
<th>x=0</th>
<th>x=1</th>
</tr>
</thead>
<tbody>
<tr>
<td>y=0</td>
<td>0</td>
</tr>
<tr>
<td>y=1</td>
<td>1/2</td>
</tr>
</tbody>
</table>

"and"

"if x \& y are obs., P x \& y \& box 2 = \int \int_{box} p(x, y) dx dy"

Marginal distribution (in the context of the joint)

Distribution of a component of a random vector

$$P\{X = x\} = \sum_{y} P\{X = x, Y = y\}$$

"sum rule"

"marginalizing out y"
R.V. $X_1, \ldots, X_n$ are "mutually independent" $\iff p(x_1, \ldots, x_n) = \prod_{i=1}^{n} p(x_i) \quad \forall x_{1:n} \in \prod_{i=1}^{n} \mathbb{R}$

**Conditioning:**

- For events $A$ and $B$, suppose $P(B) \neq 0$
  
  then define $P(A|B) \triangleq \frac{P(A \cap B)}{P(B)}$

- For a discrete R.V.
  
  define "conditional pmf"
  
  $p(x|y) \triangleq P(X=x|Y=y) \triangleq P(\{X=x, Y=y\}) \triangleq \frac{P(\{X=x, Y=y\})}{\sum_{y} P(\{Y=y\})}$

- For cts. random variable
  
  "conditional pdf"
  
  $p(x|y) \triangleq \frac{p(x,y)}{p(y)}$ density

  subtle point: $(p_2|y)$ is undefined when $p(y) = 0$
Example: \[ P(\text{having cancer} | \text{tumor measurement} = a) \]

\[ cdf \quad F_X(z) = P(X \leq z) \]