Probability 2

MML 6.2, 6.3

Marek Petrik

8/30/2023

Probability Space

- **Outcome space**: Set Ω

  Finite or infinite

  $\Omega = \{J, M, E\}$ $\Omega = [0,1]$

- **Event space** (sigma algebra): $\mathcal{F} \subset 2^{\Omega}$

  $\mathcal{F} = \{ \{J, M, E\}, \emptyset, \{J\}, \ldots \}$

- **Probability function**: $P\colon \mathcal{F} \to [0,1]$

  $P(\Omega) = 1$

  $P(\{J\}) = 0.5$

Random Variable

A $\mathcal{T}$-valued random variable X (upper case!) is a function:

$$X\colon \Omega \to \mathcal{T}$$

Examples $\Omega = \{J, E, M\}$:

$H(J) = 1$
$H(E) = 2$
$\vdots$

Random Variable: Pre-image (inverse)

Pre-image $X^{-1}\colon \mathcal{T} \to 2^{\Omega}$ defined as

$$X^{-1}(x) = \{\omega \in \Omega \mid X(\omega) = x\}$$

Examples $\Omega = \{J, E, M\}$:

$H(J) = 1$ $H(E) = 2$ $H(M) = 1$

$H^{-1}(1) = \{J, M\}$

$H^{-1}(2) = \{E\}$

$H^{-1}(3) = \emptyset$

Probability Distribution

Always associated with a random variable for some $X : \Omega \to \mathcal{T}$

$$\mathbb{P}[X = x] = P\left(X^{-1}(x)\right) = P(\{\omega \in \Omega \mid X(\omega) = x\})$$

$\mathbb{P}[H = 1] = P(\{J\}) = 0.5$

Today

1. Discrete probability distribution
2. Joint distribution
3. Sum rule, product rule
4. Bayes theorem
5. Continuous probability distribution
6. Expectation(maybe)

Joint Probability

Discrete random variables X, Y

AND　　　　　　　　　　　　intersection

$$\mathbb{P}[X = x, Y = y] = P(X^{-1}(x) \cap Y^{-1}(y)) = \dots$$

Example: $P[H = 1, W = 3]$

Probability mass function: $p_{X,Y}(x, y) = \mathbb{P}[X = x, Y = y]$

$P(X^{-1}(x) \wedge X^{-1}(y)) = P(\{\omega \in \Omega \mid X(\omega) = x\} \cap \{\omega \in \Omega \mid Y(\omega) = y) =$

$= P(\omega \in \Omega \mid X(\omega) = x \wedge Y(\omega) = y)$

Y

	x_1	x_2	x_3
y_1	0.1	0.1	0.1
y_2	0.1	0	0.1
y_3	0	0.1	0
y_4	0.2	0	0.1

X

$$\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_{X,Y}(x, y) = 1$$

Marginalization (Sum Rule)

Know $\mathbb{P}[X = x, Y = y]$ and need to compute $\mathbb{P}[X = x]$:

$$\mathbb{P}[X = x] = \sum_{y \in \mathcal{Y}} \mathbb{P}[X = x, Y = y]$$

$\mathcal{Y}$ ← domain of Y

Y		x_1	x_2	x_3
	y_1	0.2	0.1	0.2
	y_2	0.4	0.1	0.0

X

$\mathbb{P}[Y = y_1] =$

$\mathbb{P}[X = x_2] =$

Conditional Probability

Random variables: $X, Y : \Omega \to \mathcal{T}$ <u>Definition</u>:

$$\mathbb{P}[X = x \mid Y = y] = \frac{\mathbb{P}[X = x, Y = y]}{\mathbb{P}[Y = y]}$$

<u>Note</u>: $\sum_{x \in \mathcal{X}} \mathbb{P}[X = x \mid Y = y] = 1$ (when defined)

Y \ X	x_1	x_2	x_3
y_1	0.2	0.3	0.0
y_2	0.4	0.1	0.0

$\mathbb{P}[X = x_1 \mid Y = y_1] =$

$\mathbb{P}[Y = y_2 \mid X = x_2] =$

$\mathbb{P}[Y = y_2 \mid Y = y_2] = 1$

$\mathbb{P}[Y = y_1 \mid X = x_3] =$

$\mathbb{P}[X = x \mid Y = y]$ is undefined when $\mathbb{P}[Y = y] = 0$

See Borel-Kolmorogov paradox

Product Rule

Random variables: $X, Y : \Omega \to \mathcal{T}$

$$\mathbb{P}[X = x, Y = y] = \mathbb{P}[X = x \mid Y = y] \cdot \mathbb{P}[Y = y]$$

- Reverse of conditional probability

- Marginalization can then be expressed as:

$$\mathbb{P}[X=x] = \sum_{y \in \mathcal{Y}} \mathbb{P}[X=x, Y=y] = \sum_{y \in \mathcal{Y}} \mathbb{P}[X=x \mid Y=y]\, \mathbb{P}[Y=y]$$

Independence

Random variables $X, Y : \Omega \to \mathbb{R}$ are **independent** if for all $x, y \in \mathbb{R}$:

$$\mathbb{P}[X = x, Y = y] = \mathbb{P}[X = x] \cdot \mathbb{P}[Y = y]$$

Examples?

X, Y are independent

$$\mathbb{P}[X = x \mid Y = y] = \frac{\mathbb{P}[X = x, Y = y]}{\mathbb{P}[Y = y]} = \frac{\mathbb{P}[X = x]\,\mathbb{P}[Y = y]}{\mathbb{P}[Y = y]} =$$

$$= \mathbb{P}[X = x]$$

Intuitive meaning: knowing Y gives you no information about X.

- Two coin flips
- Two die rolls
- Two errors in sensor readings

$p_{X,Y}$ as a matrix is rank 1

Bayes Theorem

Invert conditional probabilities

Random variables: $X, Y : \Omega \to \mathcal{T}$

$$\mathbb{P}[X = x \mid Y = y] = \frac{\mathbb{P}[Y = y \mid X = x] \cdot \mathbb{P}[X = x]}{\mathbb{P}[Y = y]}$$

Express $\mathbb{P}[X=x \mid Y=y]$ using $\mathbb{P}[Y=y \mid X=x]$

Proof:

$$\mathbb{P}[X=x \mid Y=y] \underset{\substack{\uparrow \\ \text{definition}}}{=} \frac{\mathbb{P}[X=x, Y=y]}{\mathbb{P}[Y=y]} \underset{\substack{\uparrow \\ \text{product rule}}}{=} \frac{\mathbb{P}[Y=y \mid X=x]\,\mathbb{P}[X=x]}{\mathbb{P}[Y=y]}$$

Also a useful form:

$$\mathbb{P}[X=x, Y=y] = \frac{\mathbb{P}[Y=y \mid X=x]\,\mathbb{P}[X=x]}{\sum_{x' \in \mathcal{X}} \mathbb{P}[Y=y \mid X=x']\,\mathbb{P}[X=x']}$$

Bayes Theorem: Inverting Probabilities

$$\mathbb{P}[X = x \mid Y = y] = \frac{\mathbb{P}[Y = y \mid X = x] \cdot \mathbb{P}[X = x]}{\mathbb{P}[Y = y]}$$

Prevalence: 1%, Pos. when sick: 90%, Neg. when healthy: 95%

Probability of sick when positive test:

$S: \Omega \to \{0,1\}$ $\quad T: \Omega \to \{0,1\}$

$\mathbb{P}[S=1] = 0.01$

$\mathbb{P}[S=0] = 0.99$

$\mathbb{P}[T=1 \mid S=1] = 0.9$

$\mathbb{P}[T=0 \mid S=1] = 0.1$

$\mathbb{P}[T=1 \mid S=0] = 0.05$

$\mathbb{P}[T=0 \mid S=0] = 0.95$

solution:

$$\mathbb{P}[S=1 \mid T=1] = \frac{\mathbb{P}[T=1 \mid S=1]\,\mathbb{P}[S=1]}{\mathbb{P}[T=1]}$$

$$\mathbb{P}[T=1] = \mathbb{P}[T=1 \mid S=1]\,\mathbb{P}[S=1] + \mathbb{P}[T=1 \mid S=0]\,\mathbb{P}[S=0]$$

$$= 0.9 \cdot 0.01 + 0.05 \cdot 0.99 =$$

$$\mathbb{P}[T=1 \mid S=1]\,\mathbb{P}[S=1] = 0.9 \cdot 0.01 =$$

$$\mathbb{P}[S=1 \mid T=1] =$$

Discrete Probability Distributions

- **Bernoulli**: Heads or tails
- **Binomial**: Number of heads
- **Geometric**: Coin flips until heads
- **Poisson**: Number of customers

See Wikipedia for their properties

Continuous Random Variable

Usually real-valued: $X : \Omega \to \mathbb{R}$

Example: Probability space

- $\Omega = [0, 1]$ ← Sufficient, no need to use all $\mathbb{R}$
- $\mathcal{F}$ = Borel σ-algebra of all intervals (open and closed) Cannot use $2^{\mathbb{R}}$
- $P([a, b]) = b - a$ for $a \leq b$ (Lebesgue mesure)

Random variable: $X(\omega) = \omega$ for each $\omega \in \Omega$ Uniform random variable

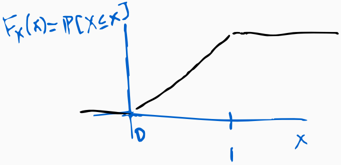
$$\mathbb{P}[X = 0.5] = P([0.5, 0.5])\ ?$$

$$= 0$$

Cumulative Distribution Function

$$F_X(x) \; = \; \mathbb{P}[X \leq x] = P(\{\omega \in \Omega \mid X(\omega) \leq x\})$$

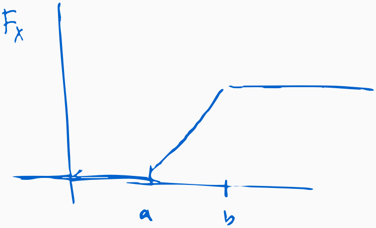
Example: Uniform random variable on $[0, 1]$

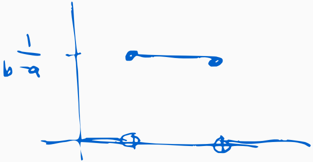


Function $f_X : \mathbb{R} \to \mathbb{R}$ is a pdf of X if $\forall x$

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

If exists f_X is the derivative of F_X



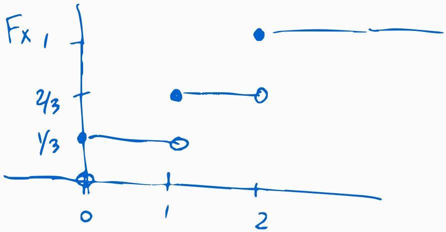


Atomic Random Variables

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

Random variable X need not have pdf!

Let $\mathbb{P}[X = 0] = \mathbb{P}[X = 1] = \mathbb{P}[X = 2] = \frac{1}{3}$



Continuous Probability Distributions

- **Normal**: common because of central limit theorem
- **Laplace**: Extreme weather events
- **Multivariate normal**: Height and weight

See Wikipedia for their properties

Expected Value (Mean)

Random variable X (discrete and continuous)

$$\mathbb{E}[X] = \sum_{x \in \mathcal{X}} x \cdot \mathbb{P}[X = x], \qquad \mathbb{E}[X] = \int_{\Omega} X dP$$

Expected Value of Function

Random variable X (discrete and continuous) and a function g

$$\mathbb{E}\left[g(X)\right] = \sum_{g \in \mathcal{G}} g \cdot \mathbb{P}\left[g(X) = g\right], \qquad \mathbb{E}\left[X\right] = \int_{\Omega} g(X) dP$$

Law of the Unconscious Statistician proves that: