

Discrete probability measures

- ▶ Consider a chance experiment with outcomes in a set Ω where Ω is finite or countably infinite, i.e., Ω is of the form

$$\Omega = \{\omega_1, \dots, \omega_n\} \quad \text{or} \quad \Omega = \{\omega_1, \omega_2, \dots\}.$$

- ▶ The probabilities are given by a real-valued function

$$\text{Prob} : \Omega \rightarrow [0, 1]$$

such that the values $\text{Prob}[\omega]$ add up to 1.

Prob is a *discrete probability distribution*, (Ω, Prob) is a *discrete probability space*.

- ▶ A subset of Ω is called an *event*. The function Prob can be extended to all events E by letting

$$\text{Prob}[E] = \sum_{\omega \in E} \text{Prob}[\omega].$$

Uniform measure

- ▶ The *uniform measure* on a finite set Ω is given by $\text{Prob}[\omega] = \frac{1}{|\Omega|}$ for all $\omega \in \Omega$.
E.g., a cast of a fair die can be modelled by the *uniform distribution* on $\Omega = \{1, \dots, 6\}$.
- ▶ On a countably infinite set there is no uniform measure.

Random variables

- ▶ A *random variable* is a mapping $X : \Omega \rightarrow \mathbb{R}$.
- ▶ Any random variable X that is defined on a discrete probability space (Ω, Prob) defines a discrete probability measure Prob_X on its range $\Omega_X = \{X(\omega) : \omega \in \Omega\}$ where

$$\text{Prob}_X[x] = \sum_{\{\omega \in \Omega : X(\omega) = x\}} \text{Prob}[\omega].$$

Prob_X is called the *distribution* of X .

- ▶ We write $\text{Prob}[X = x]$ instead of $\text{Prob}_X[x]$, and also use notation such as $\text{Prob}[X \geq x]$, $\text{Prob}[X \in S]$, \dots .

Indicator variables

- ▶ The *indicator variable* for a set $A \subseteq \Omega$ is the random variable $X : \Omega \rightarrow \{0, 1\}$ such that $X(\omega) = 1$ holds if and only if $\omega \in A$.

Example: $\Omega = \{1, \dots, 6\}$, $\text{Prob}[\omega] = \frac{1}{6}$ for all $\omega \in \Omega$.

Consider the indicator variable $X : \Omega \rightarrow \mathbb{R}$ for the set of primes less than or equal to 6

$$X(i) = \begin{cases} 1 & \text{in case } i \text{ is prim} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Prob}[X = 0] = \text{Prob}[\{1, 4, 6\}] = 1/2,$$

$$\text{Prob}[X = 1] = \text{Prob}[\{2, 3, 5\}] = 1/2.$$

Joint distribution

- ▶ Let X_1, \dots, X_m be random variables on the same discrete probability space Ω .
- ▶ The *joint distribution* $\text{Prob}_{X_1, \dots, X_m}$ is

$$\text{Prob}_{X_1, \dots, X_m}[r_1, \dots, r_m] = \sum_{\{\omega \in \Omega: X_1(\omega)=r_1, \dots, X_m(\omega)=r_m\}} \text{Prob}[\omega]. \quad (1)$$

- ▶ We write $\text{Prob}[X_1 = r_1, \dots, X_m = r_m]$ instead of $\text{Prob}_{X_1, \dots, X_m}[r_1, \dots, r_m]$.

Joint distribution

Example: $\Omega = \{1, \dots, 6\}$, $\text{Prob}[\omega] = \frac{1}{6}$ for all $\omega \in \Omega$.

Consider indicator variables X , Y and Z for the events ω is prime, ω is even, and ω is odd.

X , Y , and Z have the same distribution, the uniform distribution on $\{0, 1\}$. However,

$$\begin{aligned} \text{Prob}[X = 1, Y = 1] &= \frac{1}{6}, \\ \text{Prob}[X = 1, Z = 1] &= \frac{2}{6}. \end{aligned}$$

So the joint distribution is not determined by the individual distributions.

Mutually independent random variables

- ▶ Let X_1, \dots, X_m be random variables on the same discrete probability space Ω .
- ▶ X_1, \dots, X_m are *mutually independent* if for any combination of values r_1, \dots, r_m in the range of X_1, \dots, X_m , respectively,

$$\begin{aligned} \text{Prob}[X_1 = r_1, \dots, X_m = r_m] \\ = \text{Prob}[X_1 = r_1] \cdot \dots \cdot \text{Prob}[X_m = r_m]. \end{aligned}$$

Example: m tosses of a fair coin

Consider m tosses of a fair coin and let X_i be the indicator variable for the event that the i th toss shows head. Then the X_1, \dots, X_m are mutually independent and for all $(r_1 \dots r_m) \in \{0, 1\}^m$

$$\text{Prob}[X_1 = r_1, \dots, X_m = r_m] = \frac{1}{2^m}.$$

Pairwise and k -wise independence

- ▶ Random variables X_1, \dots, X_m are *pairwise independent* if all pairs X_i and X_j with $i \neq j$ are mutually independent, i.e., for all $i \neq j$ and all r_i and r_j

$$\begin{aligned} \text{Prob}[X_i = r_i \text{ and } X_j = r_j] \\ = \text{Prob}[X_i = r_i] \cdot \text{Prob}[X_j = r_j]. \end{aligned}$$

- ▶ The concept of k -wise independence of random variables for any $k \geq 2$ is defined similar to pairwise independence, where now every subset of k distinct random variables must be mutually independent.

Mutual versus pairwise independence

- ▶ It can be shown that mutual independence implies pairwise independence.
- ▶ For three or more random variables, in general pairwise independence does not imply mutual independence.
- ▶ It can be shown for any $k \geq 2$ that $(k + 1)$ -wise independence implies k -wise independence, whereas the reverse implication is false.
- ▶ Examples of random variables that are k -wise independent but are not mutually independent will be constructed in the section on derandomization.
An even simpler example is the following.

Pairwise and 3-wise independence

Example: pairwise but not 3-wise independence.

Consider a chance experiment where a fair coin is tossed 3 times and let X_i be the indicator variable for the event that coin i shows head. Let

$$Z_1 = X_1 \oplus X_2, \quad Z_2 = X_1 \oplus X_3, \quad Z_3 = X_2 \oplus X_3 .$$

The random variables Z_1 , Z_2 , and Z_3 are pairwise independent because for any pair i and j of distinct indices in $\{1, 2, 3\}$ and any values b_1 and b_2 in $\{0, 1\}$ we have

$$\text{Prob}[Z_i = b_1 \& Z_j = b_2] = 1/4 .$$

On the other hand, Z_1 , Z_2 , and Z_3 are not 3-wise independent because for example we have $Z_1 = Z_2 \oplus Z_3$.

Expectation

- ▶ The *expectation* of X is

$$\mathbf{E}[X] = \sum_{\omega \in \Omega} \text{Prob}[\omega] X(\omega) ,$$

provided that this sum converges absolutely. If the latter condition is satisfied, we say the expectation of X exists.

- ▶ Recall that
 - ▶ $\sum_{i \in \mathbb{N}} a_i$ converges to s if and only if the partial sums $a_0 + \dots + a_n$ converge to s ,
 - ▶ $\sum_{i \in \mathbb{N}} a_i$ converges absolutely if even the sum $\sum_{i \in \mathbb{N}} |a_i|$ converges,
 - ▶ absolute convergence is equivalent to convergence to the same value under arbitrary reorderings.
- ▶ The condition on absolute convergence ensures that the expectation is the same no matter how we order Ω .
The condition is always satisfied if Ω is finite or if X is non-negative and the sum converges at all.

Expectation

Example: $\Omega = \{1, \dots, 6\}$, $\text{Prob}[\omega] = \frac{1}{6}$ for all $\omega \in \Omega$.

If we let X be the identity mapping on Ω , then

$$\mathbf{E}[X] = \sum_{i \in \{1, \dots, 6\}} \text{Prob}[i] X(i) = \frac{1}{6} + \frac{2}{6} + \dots + \frac{6}{6} = \frac{21}{6} = 3.5 .$$

Example: $\Omega = \mathbb{N}$, $\text{Prob}[i] = \frac{1}{2^{i+1}}$.

The expectation of the random variable

$$X : i \mapsto 2^{i+1}$$

does not exist because the corresponding sum does not converge

$$\sum_{i \in \mathbb{N}} \text{Prob}[i] X(i) = \sum_{i \in \mathbb{N}} \frac{1}{2^{i+1}} 2^{i+1} = 1 + 1 + \dots = +\infty .$$

Linearity of Expectation

- ▶ Let X and X_1, \dots, X_n be random variables such that their expectations all exist.
- ▶ Expectation is linear.

For any real number r , the expectation of rX exists and it holds that

$$\mathbf{E}[rX] = r\mathbf{E}[X].$$

The expectation of $X_1 + \dots + X_m$ exists and it holds that

$$\mathbf{E}[X_1 + \dots + X_n] = \mathbf{E}[X_1] + \dots + \mathbf{E}[X_n].$$

- ▶ If the X_1, \dots, X_n are *mutually independent*, then the expectation of $X_1 \cdot \dots \cdot X_m$ exists and

$$\mathbf{E}[X_1 \cdot \dots \cdot X_n] = \mathbf{E}[X_1] \cdot \dots \cdot \mathbf{E}[X_n].$$

Conditional distributions and expectations

- ▶ The *conditional probability* of an event E given an event F is

$$\text{Prob}[E|F] = \frac{\text{Prob}[E \cap F]}{\text{Prob}[F]},$$

where this value is undefined in case $\text{Prob}[F] = 0$.

- ▶ The *conditional distribution* $\text{Prob}[\cdot|F]$ of a random variable X given an event F is defined by

$$\text{Prob}[X = a|F] = \frac{\text{Prob}[\{\omega \in \Omega : X(\omega) = a\} \cap F]}{\text{Prob}[F]}.$$

- ▶ The *conditional expectation* $\mathbf{E}[X|F]$ of a random variable X given an event F is the expectation of X with respect to the conditional distribution $\text{Prob}[X|F]$, i.e.,

$$\mathbf{E}[X|F] = \sum_{a \in \text{range}(X)} a \cdot \text{Prob}[X = a|F].$$

Number of fixed points of a random permutation

- ▶ Suppose that n tokens T_1, \dots, T_n are distributed at random among n persons P_1, \dots, P_n such that each person gets exactly one token and all such assignments of tokens to persons have the same probability (i.e., the tokens are assigned by choosing a permutation of $\{1, \dots, n\}$ uniformly at random).
- ▶ What is the expected number of indices i such that P_i gets his or her "own token" T_i ?
If we let X_i be the indicator variable for the event that P_i gets T_i , then $X = \sum_{i=1}^n X_i$ is equal to the random number of persons that get their own token.
- ▶ By linearity of expectation, the expectation of X is

$$\mathbf{E}[X] = \mathbf{E}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \mathbf{E}[X_i] = \sum_{i=1}^n \left(\frac{n-1}{n} \cdot 0 + \frac{1}{n} \cdot 1\right) = 1.$$

Conditional distributions and expectations

Markov Inequality

Proposition (Markov Inequality)

Let X be a random variable that assumes only non-negative values. Then for every positive real number r , we have

$$\text{Prob}[X \geq r] \leq \frac{\mathbf{E}[X]}{r}.$$

Proof.

Let (Ω, Prob) be the probability space on which X is defined. Then we have

$$\begin{aligned} \mathbf{E}[X] &= \sum_{\omega \in \Omega} \text{Prob}[\omega] X(\omega) \geq \sum_{\{\omega \in \Omega : X(\omega) \geq r\}} \text{Prob}[\omega] X(\omega) \\ &\geq r \sum_{\{\omega \in \Omega : X(\omega) \geq r\}} \text{Prob}[\omega] \geq r \text{Prob}[X \geq r]. \end{aligned}$$

□