Toronto Metropolitan University, Department of Computer Science.
Reinforcement Learning: Probability Formula Sheet

1 Conditional Probability

- Consider running a probabilistic experiment, e.g., throwing a six-sided die or tossing a coin. The set of all possible outcomes of the experiment is called the *sample space*. If the six-sided die is fair, each outcome has equal probability $\frac{1}{6}$. A (random) *event* is a subset of the sample space, e.g., an event E of throwing an even number $\{2, 4, 6\}$ with a fair die, or an event O of throwing an odd number $\{1, 3, 5\}$. These are two mutually exclusive events, since $E \cap O$ is empty and $E \cup O$ is the whole sample space.
- Let V be an event with non-zero probability, i.e., $P(V) > 0$. $P(\, E \,|\, V \,)$ is the probability of the occurence of event E given that V occured and is given as

$$P(\, E \,|\, V \,) = \frac{P(E \cap V)}{P(V)}$$

  Knowing that V occured reduces the sample space to V, and the part of it where E also occured is $E \cap V$. A common abbreviation for probability $P(E \cap V)$ is either $P(E\,V)$, or $P(E, V)$.
- Because $\cap$ is commutative, we have $P(E \cap V) = P(E \,|\, V) \cdot P(V) = P(V \,|\, E) \cdot P(E)$.
- When random events V_i are mutually exclusive and exhaustive set of events (i.e., events V_i and V_j cannot co-occur if $i \neq j$ and $\bigcup_i V_i$ is the whole sample space), then

$$P(E) = \sum_i P(E \cap V_i) = \sum_i P(E|V_i) \cdot P(V_i)$$

- If random events E and V are independent, then $P(E|V) = P(E)$ and hence $P(E \cap V) = P(E) \cdot P(V)$ That is, knowledge that whether an event V has occured does not change the probability that E occurs.

2 A Cumulative Distribution Function

- A *random variable* (r.v.) is a function that maps an outcome of a probabilistic experiment (a chance) to a value. If a random variable takes on a finite number of values, then it is called *discrete*. Otherwise, it is *continuous*. For example, "the number of heads in two tosses of a coin" is a discrete random variable.
- A cumulative distribution function (cdf) $F(\cdot)$ of a random variable X for any real number a is

$$F(a) = P\{X \leqslant a\}$$

  If a cdf $F(\cdot)$ is given, we calculate probability of X falling in the range between a and b as

$$P\{a < X \leqslant b\} = F(b) - F(a)$$

- If X is a discreet random variable then

$$F(a) = \sum_{\forall x \leqslant a} P(x)$$

  where $P(\cdot)$ is the function defined as $P(x) = P\{X = x\}$, and $P\{X = x\}$ means the probability of the event that a variable X takes the value x in a random experiment.

3 Properties of Expectation. Conditional Expectation

- Expectation, or expected value, or mean of a discrete random variable X, denoted by $E[X]$, is the sum of the values that X can take multiplied by the probabilities that each value can occur

$$E[X] = \sum_i x_i \cdot P(x_i)$$

where x_i is one of the possible values of X and $P(x_i) = P\{X = x_i\}$ is the probability of the event that the variable X takes the value x_i in a random experiment. In other words, it is a weighted average where each value is weighted by the probability that X takes that value.

- It is a well-known result that if a and b are real numbers, X and Y are random variables, then no matter if X, Y are independent or not

$$E[a \cdot X + b] = a \cdot E[X] + b \quad (1)$$
$$E[X + Y] = E[X] + E[Y] \quad (2)$$

- The conditional expectation of a discrete random variable X given that a random variable Y takes the value y, i.e. $Y = y$, is defined by

$$E[X \mid Y = y] = \sum_i x_i \cdot P\{X = x_i \mid Y = y\}$$

- Consider a new random variable $E[X|Y]$ that takes values $E[X|Y = y_j]$ with probability $P\{Y = y_j\}$. Because the conditional expectation $E[X|Y]$ is a random variable, if Y is a discrete random variable, then by the definition of expectation we can compute the expected value of $E[X|Y]$ as

$$E[\,E[X|Y]\,] = \sum_j E[X|Y = y_j] \cdot P\{Y = y_j\}.$$

In reinforcement learning, a r.v. $E[X|Y]$ is important because a random reward is conditioned on the previous state and an action taken by the agent. Exercise: show a standard result that

$$E[X] = E[\,E[X|Y]\,] = \sum_j E[X|Y = y_j] \cdot P\{Y = y_j\}$$

This equation states that to calculate $E[X]$, we can take a weighted average of the conditional expected value of X given that $Y = y_j$, each of the terms $E[X|Y = y_j]$ being weighted by the probability $P\{Y = y_j\}$ of the event on which it is conditioned.

4 A Law of Large Numbers

The mean (the expected value) can be approximated by the sample average. Recall that the variance σ^2 of X is the expectation of the squared deviation of a random variable X from its mean, $E[(X - E[X])^2]$. Let $\{X_t\}_{t=1}^N$ be a set of N independent and identically distributed (i.i.d.) random variables each having mean μ and a finite variance σ^2. Then for any $\epsilon > 0$

$$P\left\{ \left| \frac{\sum_t X_t}{N} - \mu \right| > \epsilon \right\} \to 0 \qquad as \quad N \to \infty$$

That is, the average value of X_t over N trials converges to the mean value of X_t as N increases.

5 Well-known Continuous Random Variables

5.1 Uniform Distribution

- A random variable X is uniformly distributed over the interval [a, b] if its density function is given by

$$p(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leqslant x \leqslant b \\ 0 & \text{otherwise} \end{cases}$$

- If X is uniform, its expected value and variance σ^2 are

$$E[X] = \frac{a+b}{2}, \qquad \sigma^2(X) = \frac{(b-a)^2}{12}$$

5.2 Normal (Gaussian) Distribution

- A random variable X is normally distributed with mean μ and variance σ^2, denoted as $N(\mu, \sigma^2)$, if its density function is

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad -\infty < x < \infty$$

6 Generating Random Numbers in C

Read the manual pages "man *random*" and "man *srand*" on any Linux computer for additional details.

- The following C library functions can be used to generate a random value according to a uniform distribution between 0 and 1:

```
#include <time.h>     /* to declare function time() */
#include <stdlib.h>   /* to declare functions rand() and srand() */
/* call srand once with the current time to seed random generator*/
srand( (unsigned int) time(NULL));
/* rand() returns a random integer between 0 and RAND_MAX*/
int r = rand();
```

- To generate a random number according to the Gausian distribution, the GNU Scientific Library can be used. For information about how to use this library, about Unix random generators, or about the Gaussian random functions, check out these links

  https://www.gnu.org/software/gsl/doc/html/randist.html#the-gaussian-distribution

  https://www.gnu.org/software/gsl/doc/html/rng.html

  Example: the function *double gsl_cdf_gaussian_P (double a, double sigma)* calculates a Gaussian cdf as in Section 2, where a is a parameter, and *sigma* is standard deviation which is $\sqrt{\sigma^2}$ of variance σ^2 mentioned above. This function returns probability that a Gaussian random variable is $\leq a$.

7 Generating Random Numbers in Java

Information related to random number generating in Java can be found at
`http://docs.oracle.com/javase/7/docs/api/java/util/Random.html`
`https://docs.oracle.com/en/java/javase/19/docs/api/java.base/java/util/Random.html`