Lectures 23-24

Andrei Antonenko

March 18, March 28, 2005

1 Stochastic processes

Stochastic process is the model of the experiment which evolves in time, and generates the sequence of numbers. For example, the sequence of stock prices, sequence of time of device failure, sequence of the scores in the football game can all be viewed as stochastic processes. The first type of the process we will concentrate on is a Bernoulli Process.

2 Bernoulli process

Bernoulli process can be viewed as a sequence of independent coin tosses with probability of heads equal to p (obviously, we use coin only for presentation purposes – normally, there are experiments which has either success of failure as outcome, and success happens with probability p). Formally, Bernoulli process is a sequence of random variables $X_1, X_2, \dots$, such that

$$P(X_i = 1) = P(\text{success on } i\text{-th trial}) = p$$
$$P(X_i = 0) = P(\text{failure on } i\text{-th trial}) = 1 - p$$

What kind of question and which random variables we can be interested in such a process? A lot of facts are already known to us from the previous lectures.

- S – the number of successes in n independent trials. The distribution of S is binomial with parameters n and p:

$$p_S(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots$$

$$\mathbf{E}[S] = np, \qquad \mathbf{var}(S) = np(1-p).$$

- Number of trials up to the first success T. This is a geometrically distributed random variable:

$$p_T(t) = (1-p)^{t-1} p, \quad t = 1, 2, \dots$$
$$\mathbf{E}[T] = \frac{1}{p}, \qquad \mathbf{var}(T) = \frac{1-p}{p^2}$$

2.1 Independence, memorylessness

Now, all the trials are independent. Therefore, for example, the number of successes in the first 5 trials $U = X_1 + X_2 + X_3 + X_4 + X_5$ is independent of the number of successes in the second 5 trials $V = X_6 + X_7 + X_8 + X_9 + X_{10}$. Another example we can consider is the following. Let U (V) be the first odd (even) time at which there was a success. Since U is determined by the sequence $X_1, X_3, \ldots,$ and V is determined by the sequence $X_2, X_4, \ldots,$ they are independent, since these sequences don't have common elements.

Now, assume we observed the Bernoulli process up to stage n, i.e. we observed the random variables $X_1, \ldots, X_n$. Now, the future trials of the process $X_{n+1}, X_{n+2}, \ldots$ are independent of the past ones. Therefore, in the Bernoulli process the future of it is independent from the past.

Now let's recall that T is a time till the first success. Assume we are watching the process for n stages already. What can be said about the number $T - n$ of remaining trials until the first success? We have:

$$\begin{aligned}
P(T-n=t|T>n) &= \frac{P(T-n=t, T>n)}{P(T>n)} \\
&= \frac{P(T=t+n)}{P(T>n)} \\
&= \frac{p(1-p)^{t+n-1}}{1-P(T\leq n)} \\
&= \frac{p(1-p)^{t+n-1}}{1-p\frac{1-(1-p)^n}{1-(1-p)}} \\
&= \frac{p(1-p)^{t+n-1}}{(1-p)^n} \\
&= p(1-p)^{t-1} \\
&= P(T=t).
\end{aligned}$$

That means that the number of future trials until the first success is described by the same geometric PMF with parameter p and is independent of the past. This is called a **memoryless property** of Bernoulli process.

2.2 Interarrival times

An important random variable associated with the Bernoulli process is the time of the k-th success, which we denote by Y_k. A related random variable is the k-th interarrival time, denoted by T_k. It is defined by

$$T_1 = Y_1, T_k = Y_k - Y_{k-1}, \quad k = 2, 3, \ldots \tag{1}$$

and represents the number of trials between $k-1$-st and k-th successes. Moreover, we can note that

$$Y_k = T_1 + T_2 + \cdots + T_k.$$

The time T_1 until the first success is a geometric random variable with parameter p. After the success, the future of the process does not depend on the past, and therefore, the time until the second success will not be dependent on the past of the process, and will still be geometrically distributed with parameter p. Therefore, we deduce, that all T_i's are independent geometrically distributed random variables with parameter p.

2.3 k-th arrival time

The time Y_k of the k-th success is equal to the sum of k independently distributed geometric random variables:

$$Y_k = T_1 + T_2 + \cdots + T_k.$$

That allows us to get the formulas for expectation and variance of Y_k:

$$\mathbf{E}[Y_k] = \mathbf{E}[T_1] + \cdots + \mathbf{E}[T_k] = \frac{k}{p}$$
$$\mathbf{var}(Y_k) = \mathbf{var}(T_1) + \cdots + \mathbf{var}(T_k) = \frac{k(1-p)}{p^2}$$

Now let's derive the formula for the PMF of Y_k. First let's notice that $Y_k \geq k$. Now, for $t \geq k$ the event $Y_k = t$ occurs if both of the following conditions are true:

(i) event A: trial t is a success;

(ii) event B: exactly $k-1$ successes happened in the first $t-1$ trials.

We have:

$$P(A) = p$$
$$P(B) = \binom{t-1}{k-1} p^{k-1}(1-p)^{t-k}.$$

Moreover, the events A and B are independent, since the t-th trial does not depend on the past – trials from 1 to $t-1$. Therefore,

$$P(Y_k = t) = P(A \cap B) = P(A)P(B) = \binom{t-1}{k-1} p^k (1-p)^{t-k} \tag{2}$$

This PMF is called the **Pascal PMF of order** k.

2.4 Splitting and Merging of Bernoulli Processes

Starting with a Bernoulli process in which there is a probability p of success at each time, consider splitting it as follows. Whenever there is a success, we choose to either keep it (with probability q), or to discard it (with probability $1-q$).

Assume that the decisions to keep or discard are independent for different successes. If we focus on the process of successes that are kept, we see that it is a Bernoulli process: in each time slot, there is a probability pq of a kept success, independently of what happens in other slots. For the same reason, the process of discarded successes is also a Bernoulli process, with a probability of a discarded arrival at each time slot equal to $p(1-q)$.

In a reverse situation, we start with two independent Bernoulli processes (with parameters p and q, respectively) and merge them into a single process, as follows. A success is recorded in the merged process if and only if there is a success in at least one of the two original processes, which happens with probability $p + q - pq$ (one minus the probability $(1-p)(1-q)$ of no success in either process.) Since different time slots in either of the original processes are independent, different slots in the merged process are also independent. Thus, the merged process is Bernoulli, with success probability $p + q - pq$ at each time step.

2.5 Poisson approximation to Binomial

Number of successes in n trials is a Binomial random variable with parameters n and p and the means np. Now we will concentrate on the case when n is large and p is small. One can think of the number of accidents: the number of cars in large, but the probability of the individual car getting into the accident in small.

Mathematically speaking, we will increase the value of n and decrease the value of p, keeping their product constant: $\lambda = np$. In the limit, it turns out that the formula for the binomial PMF simplifies to the Poisson PMF.

Proof. Setting $p = \lambda/n$ we have:

$$\begin{aligned} p_S(k) &= \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k} \\ &= \frac{n(n-1)\ldots(n-k+1)}{k!} \cdot \frac{\lambda^k}{n^k} \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{n}{n} \cdot \frac{n-1}{n} \ldots \frac{n-k+1}{n} \cdot \frac{\lambda^k}{k!} \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k}. \end{aligned}$$

Let's fix k and let $n \to \infty$. Each of the ratios $(n-1)/n, \ldots (n-k+1)/n$ converges to 1, and also

$$\left(1 - \frac{\lambda}{n}\right)^{-k} \to 1, \quad \left(1 - \frac{\lambda}{n}\right)^{n} \to e^{-\lambda}$$

Therefore,

$$p_S(k) \to e^{-\lambda} \frac{\lambda^k}{k!}.$$

□

Normally the approximation

$$e^{-\lambda} \frac{\lambda^k}{k!} \approx \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k}$$

is used when $n \geq 100, p \leq 0.01$, and $\lambda = np$.

Example 2.1. A packet consisting of a string of n symbols is transmitted over a noisy channel. Each symbol has probability $p = 0.0001$ of being transmitted in error, independently of errors in the other symbols. How small should n be in order for the probability of incorrect transmission (at least one symbol in error) to be less than 0.001?

Each symbol transmission is viewed as an independent Bernoulli trial. Thus, the probability of a positive number S of errors in the packet is $1 - P(S = 0) = 1 - (1-p)^n$. For this probability to be less than 0.001, we must have $1 - (1 - 0.0001)^n < 0.001$ or

$$n < \frac{\ln 0.999}{\ln 0.9999} = 10.0045.$$

We can also use the Poisson approximation for $P(S = 0)$, which is $e^{-\lambda}$ with $\lambda = np = 0.0001 \cdot n$, and obtain the condition $1 - e^{0.0001 \cdot n} < 0.001$, which leads to

$$n < \frac{-\ln 0.999}{0.0001} = 10.005.$$

Given that n must be integer, both methods lead to the same conclusion that n can be at most 10.