Bernoulli Process

1.1 Definition

Definition 1.1.1. A **Bernoulli process** is an infinite sequence of independent and identically (IID) distributed Bernoulli random variables X_i:

$$\mathbb{P}(X_i = x) = p^x(1-p)^{1-x}, \text{ for } x \in \{0, 1\}.$$

■

Discussion 1.1.1. Viewing each success in the Bernoulli process as an arrival, the Bernoulli process may be characterized in the following ways:

i. The sequence of arrival times (T_k), where T_k is the time of kth arrival.

ii. The sequence of inter-arrival times (Δ_k), where $\Delta_k = T_k - T_{k-1}$ is the time between the $(k-1)$st and kth arrivals.

iii. The number of arrivals by time k:

$$N_k = \max\{n \in \mathbb{N} : T_n \leq k\}.$$

♠

1.2 First Arrival Time

Proposition 1.2.1. The time to the first arrival in a Bernoulli process follows a geometric distribution:

$$\mathbb{P}(T_1 = n) = (1-p)^{n-1}p, \text{ for } n \in \mathbb{N}^{>0}$$

♦

Proof. The first arrival occurs at time n if and only if $X_1 = \cdots = X_{n-1} = 0$ and $X_n = 1$. Since each arrival indicators (X_n) are IID:

$$\mathbb{P}(X_1 = \cdots = X_{n-1} = 0, X_n = 1) = (1-p)^{n-1}p.$$

■

Proposition 1.2.2. The geometric distribution is **memory-less**:

$$\mathbb{P}(T_1 > n + m | T_1 > m) = \mathbb{P}(T_1 > n).$$

♦

Proof. The unconditional probability that the time of first arrival exceeds n is:

$$\mathbb{P}(T_1 > n) = \mathbb{P}(X_1 = \cdots = X_n = 0) = (1-p)^n.$$

The conditional probability that $T_1 > n + m$ given that $T_1 > m$ is:

$$\mathbb{P}(T_1 > n + m | T_1 > m) = \frac{\mathbb{P}(T_1 > n + m)}{\mathbb{P}(T_1 > m)} = \frac{(1-p)^{n+m}}{(1-p)^m} = (1-p)^n.$$

■

Proposition 1.2.3. The geometric distribution has constant hazard:

$$h(n) = \mathbb{P}(T_1 = n | T_1 \geq n) = p.$$

♦

Proof.

$$\mathbb{P}(T_1 = n | T_1 \geq n) = \frac{\mathbb{P}(T_1 = n)}{\mathbb{P}(T_1 \geq n)} = \frac{\mathbb{P}(T_1 = n)}{\mathbb{P}(T_1 > n-1)} = \frac{(1-p)^{n-1}p}{(1-p)^{n-1}} = p.$$

■

1.3 Arrival Times

Proposition 1.3.4. The number of arrivals by time n follows a binomial distribution:

$$\mathbb{P}(N_n = k) = \binom{n}{k} p^k (1-p)^{n-k}, \text{ for } k \in \{0, \cdots, n\}.$$

♦

Proof. Let $(x_1, \cdots, x_n)$ denote a length n binary sequence such that $\sum_{i=1}^{n} x_i = k$. Since the arrival indicators (X_i) are IID, the probability of any particular arrival sequence is:

$$\mathbb{P}\big\{(X_1, \cdots X_n) = (x_1, \cdots, x_n)\big\} = p^k(1-p)^{n-k}.$$

Now, there are $\binom{n}{k}$ distinct binary configurations $(x_1, \cdots, x_n)$ that satisfy the constraint $\sum_{i=1}^{n} x_i = k$. Therefore:

$$\mathbb{P}(N_n = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

■

Proposition 1.3.5. The time of kth arrival in a Bernoulli process follows a negative binomial distribution:

$$\mathbb{P}(T_k = n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}, \text{ for } n \geq k.$$

♦

Proof. The kth arrival occurs at time n if and only if the number of arrivals at time $n-1$ is $N_{n-1} = k-1$ and $X_n = 1$:

$$\mathbb{P}(T_k = n) = \mathbb{P}(N_{n-1} = k-1) \cdot \mathbb{P}(X_n = 1) = \binom{n-1}{k-1} p^{k-1}(1-p)^{n-k} \cdot p.$$

■

1.4 Inter-arrival Times

Discussion 1.4.1. Define $\Delta_1 = T_1$ and $\Delta_k = T_k - T_{k-1}$. The Δ_k represent **inter-arrival times**. The time of kth arrival is, by construction, equal to the sum of the first k inter-arrival times:

$$\begin{aligned} T_k &= \big(T_k - T_{k-1}\big) + \big(T_{k-1} - T_{k-2}\big) + \cdots + \big(T_2 - T_1\big) + T_1 \\ &= \Delta_k + \Delta_{k-1} + \cdots + \Delta_2 + \Delta_1 = \sum_{i=1}^{k} \Delta_i. \end{aligned}$$

♠

Proposition 1.4.6. The inter-arrival times Δ_k are independent, and each follows a geometric distribution:

$$\mathbb{P}(\Delta_k = n) = (1-p)^{n-1} p, \text{ for } n \in \mathbb{N}^{>0}.$$

♦

Proof. In terms of the arrival indicators (X_i), the kth inter-arrival time is:

$$\Delta_k = \sum_{i=T_{k-1}+1}^{T_k} X_i.$$

Since these sums are non-overlapping and the arrival indicators are IID, the Δ_k are IID in turn. Now, $\Delta_k = n$ if and only if $X_{T_k} = 1$ and:

$$X_{T_{k-1}+1} = \cdots = X_{T_k - 1} = 0$$

Thus,

$$\mathbb{P}(\Delta_k = n) = \mathbb{P}(X_{T_{k-1}+1} = \cdots = X_{T_k-1} = 0, X_{T_k} = 1) = (1-p)^{n-1} p.$$

■

Proposition 1.4.7. The sequence of arrival times (T_k) has **stationary** and **independent** increments:

i. For $k_0 < k_1$, $V_{k_1} - V_{k_0}$ has the same distribution as $V_{k_1 - k_0}$.

ii. For $k_0 < k_1 < \cdots$, the sequence of random variables:

$$(V_{k_0}, V_{k_1} - V_{k_0}, V_{k_2} - V_{k_1}, \cdots)$$

is independent.

♦

1.5 Multinomial Process

Definition 1.5.1. Suppose that each arrival in a Bernoulli process may be classified into one of J mutually exclusive types, $X_i \in \{1, \cdots, J\}$. Let $p_j = \mathbb{P}(X_i = j)$ denote the probability that an arrival is of type j. The parameters $(p_1, \cdots, p_J)$ full characterize the distribution of X_i, and satisfy the constraints:

$$p_j > 0 \text{ for } j \in \{1, \cdots, J\}, \text{ and } \sum_{j=1}^{J} p_j = 1.$$

The sequence of (X_j) is a **multinomial process**. Note that a non-arrival, coded as 0 in the Bernoulli process, may be regarded as one of the J arrival types. ■

Proposition 1.5.8. Let $N_n^{(j)}$ denote the number of type j arrivals by time n:

$$N_n^{(j)} = \sum_{i=1}^{n} \mathbb{I}(X_i = j).$$

The joint distribution of $(N_n^{(1)}, \cdots, N_n^{(J)})$ is multinomial:

$$\mathbb{P}(N_n^{(1)} = n_1, \cdots, N_n^{(J)} = n_J) = \binom{n}{n_1 \ \cdots \ n_J} p_1^{n_1} \cdots p_J^{n_J},$$

where $n = \sum_{j=1}^{J} n_j$ is the total number of arrivals. ♦

Proof. Let $(x_1, \cdots, x_n)$ denote a sequence of length n such that:

$$\sum_{i=1}^{n} \mathbb{I}(X_i = j) = n_j, \text{ for } j \in \{1, \cdots, J\}.$$

That is, the sequence $(x_1, \cdots, x_J)$ contains n_1 elements equal to 1; n_2 elements equal to 2; and so forth. Since each arrival indicators (X_i) are IID, the probability of any particular arrival sequence is:

$$\mathbb{P}\big\{(X_1, \cdots, X_n) = (x_1, \cdots, x_n)\big\} = p_1^{n_1} \cdots p_J^{n_J}.$$

Now, there are $\binom{n}{n_1 \cdots n_J}$ distinct rearrangements of the sequence $(x_1, \cdots, x_n)$ that leave the total number of arrivals of each type unchanged. Therefore:

$$\mathbb{P}(N_n^{(1)} = n_1, \cdots, N_n^{(J)} = n_J) = \binom{n}{n_1 \; \cdots \; n_J} p_1^{n_1} \cdots p_J^{n_J}.$$

■

Proof. ■

Proposition 1.5.9. The marginal distribution of the number of arrivals by time n of type j is binomial:

$$\mathbb{P}(N_n^{(j)} = n_j) = \binom{n}{n_j} p_j^{n_j}(1-p_j)^{n-n_j}, \text{ for } n_j \in \{0, \cdots, n\}.$$

♦

1.6 Beta-Bernoulli Process

Definition 1.6.1. Suppose the probability of success P is itself random, following a beta distribution with parameters a and b:

$$f(p) = \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)}, \text{ for } p \in (0,1).$$

Given P, (X_i) is a sequence of conditionally independent indicator random variables:

$$(X_i \perp X_j)|P \text{ for } i \neq j,$$

each satisfying:

$$\mathbb{P}(X_i = 1|P = p) = p.$$

Then, the sequence (X_i) is a **Beta-Bernoulli process**. ■

Proposition 1.6.10. Let N_n denote the number of arrivals by time n. The joint distribution of $(X_1, \cdots, X_n)$ given $N_n = k$ is:

$$\mathbb{P}(X_1 = x_1, \cdots, X_n = x_n | N_n = k) = \frac{B\big\{a+k, b+(n-k)\big\}}{B(a,b)}.$$

♦

Proof. By conditional independence:

$$\begin{aligned}
\mathbb{P}(X_1 = x_1, \cdots, X_n = x_n | N_n = k) &= \mathbb{E}\big\{\mathbb{P}(X_1 = x_1, \cdots, X_n = x_n | P, N_n = k) | N_n = k\big\} \\
&= \mathbb{E}\big\{p^k(1-p)^{n-k}\big\} = \frac{1}{B(a,b)} \int_0^1 p^k(1-p)^{n-k} \cdot p^{a-1}(1-p)^{b-1} dp \\
&= \frac{B\big\{a+k, b+(n-k)\big\}}{B(a,b)}.
\end{aligned}$$

■

1.7 Exercises

i. Prove that the sequence of arrival times in a Bernoulli process has stationary and independent increments.

ii. For a Bernoulli process with $m \leq n$ and $j \leq k$, argue that:

$$\mathbb{P}(N_m = j | N_n = k) = \frac{\binom{m}{j}\binom{n-m}{k-j}}{\binom{n}{k}}, \text{ for } j \in \{0, \cdots, k\}.$$

iii. Prove that the marginal distribution of $N_n^{(j)}$ in a multinomial process is binomial.

iv. Find the posterior expectation of P given $(X_1 = x_1, \cdots, X_n = x_n)$ is a Beta-Bernoulli process.