EE126: Probability and Random Processes

Lecture 17: Bernouli Process

Abhay Parekh

UC Berkeley

March 17, 2011

1. Review

2. Stochastic Processes

3. Bernoulli Process

Central Limit Theorem

Variance and mean of S_n grow unboundedly with n:
$E[S_n] = n\mu$, $var(S_n) = n\sigma^2 \Rightarrow \sigma_{S_n} = \sqrt{n}\sigma$. So why not look a the normalized version of S_n?

Central Limit Theorem

Variance and mean of S_n grow unboundedly with n:
$E[S_n] = n\mu$, $var(S_n) = n\sigma^2 \Rightarrow \sigma_{S_n} = \sqrt{n}\sigma$. So why not look a the normalized version of S_n?
Define

$$\hat{S}_n = \frac{S_n - n\mu}{\sqrt{n}\sigma}$$

($\hat{S}_n$ has mean 0 and variance 1.)

Central Limit Theorem

Variance and mean of S_n grow unboundedly with n:
$E[S_n] = n\mu$, $var(S_n) = n\sigma^2 \Rightarrow \sigma_{S_n} = \sqrt{n}\sigma$. So why not look a the normalized version of S_n?
Define

$$\hat{S}_n = \frac{S_n - n\mu}{\sqrt{n}\sigma}$$

($\hat{S}_n$ has mean 0 and variance 1.)
Then, the Central Limit Theorem says that (amazingly):

$$\boxed{\lim_{n\to\infty} P(\hat{S}_n \leq x) = \phi(x), \text{ for every } x}$$

Thus, for large n the random variable $\hat{S}_n$ approximates a standard Normal!

Central Limit Theorem

Variance and mean of S_n grow unboundedly with n:
$E[S_n] = n\mu$, $var(S_n) = n\sigma^2 \Rightarrow \sigma_{S_n} = \sqrt{n}\sigma$. So why not look a the normalized version of S_n?
Define

$$\hat{S}_n = \frac{S_n - n\mu}{\sqrt{n}\sigma}$$

($\hat{S}_n$ has mean 0 and variance 1.)
Then, the Central Limit Theorem says that (amazingly):

$$\boxed{\lim_{n\to\infty} P(\hat{S}_n \leq x) = \phi(x), \text{ for every } x}$$

Thus, for large n the random variable $\hat{S}_n$ approximates a standard Normal! I.e.

$$S_n \to N(n\mu, n\sigma^2)$$

Central Limit Theorem

Variance and mean of S_n grow unboundedly with n:
$E[S_n] = n\mu$, $var(S_n) = n\sigma^2 \Rightarrow \sigma_{S_n} = \sqrt{n}\sigma$. So why not look a the normalized version of S_n?
Define

$$\hat{S}_n = \frac{S_n - n\mu}{\sqrt{n}\sigma}$$

($\hat{S}_n$ has mean 0 and variance 1.)
Then, the Central Limit Theorem says that (amazingly):

$$\boxed{\lim_{n\to\infty} P(\hat{S}_n \leq x) = \phi(x), \text{ for every } x}$$

Thus, for large n the random variable $\hat{S}_n$ approximates a standard Normal! I.e.

$$S_n \to N(n\mu, n\sigma^2) \qquad M_n \to N(\mu, \frac{\sigma^2}{n}).$$

Implications of CLT

1. The Distributions of S_n and M_n wipe out all the information in the original information except for μ and σ^2.

Implications of CLT

1. The Distributions of S_n and M_n wipe out all the information in the original information except for μ and σ^2.
2. If there are large number of small and independent factors, the aggregate of these factors will be normally distributed. E.g. Noise.

Implications of CLT

1. The Distributions of S_n and M_n wipe out all the information in the original information except for μ and σ^2.
2. If there are large number of small and independent factors, the aggregate of these factors will be normally distributed. E.g. Noise.
3. The Gaussian Distribution is very important – many problems involve sums of iid random variables and the only thing one needs to know is the mean and variance.

Implications of CLT

1. The Distributions of S_n and M_n wipe out all the information in the original information except for μ and σ^2.
2. If there are large number of small and independent factors, the aggregate of these factors will be normally distributed. E.g. Noise.
3. The Gaussian Distribution is very important – many problems involve sums of iid random variables and the only thing one needs to know is the mean and variance.

De Moivre-Laplace Approximation to Binomial

Suppose S_n is binomial with mean np and variance $np(1-p)$.
We want to find $P(S_n = k)$.
Can't use previous example's approach since there the normal distribution is continuous....

De Moivre-Laplace Approximation to Binomial

Suppose S_n is binomial with mean np and variance $np(1-p)$.
We want to find $P(S_n = k)$.
Can't use previous example's approach since there the normal distribution is continuous....Use

$$P(S_n = k) = P(k - \frac{1}{2} \leq S_n \leq k + \frac{1}{2})$$

Now we can approximate using CLT.
Extend this idea to find $P(l \leq S_n \leq k)$ for any non-negative integers l, k.

De Moivre-Laplace Approximation to Binomial

Suppose S_n is binomial with mean np and variance $np(1-p)$.
We want to find $P(S_n = k)$.
Can't use previous example's approach since there the normal distribution is continuous....Use

$$P(S_n = k) = P(k - \frac{1}{2} \leq S_n \leq k + \frac{1}{2})$$

Now we can approximate using CLT.
Extend this idea to find $P(l \leq S_n \leq k)$ for any non-negative integers l, k.

$$P(k \leq S_n \leq l) \approx \phi(\frac{l + \frac{1}{2} - np}{\sqrt{np(1-p)}}) - \phi(\frac{k - \frac{1}{2} - np}{\sqrt{np(1-p)}})$$

Proof of the CLT

$X_1, \ldots X_n$ are iid with mean 0 and variance σ^2. Let $M_X(s)$ the transform of each of the X_i.

Proof of the CLT

$X_1, ... X_n$ are iid with mean 0 and variance σ^2. Let $M_X(s)$ the transform of each of the X_i. $\hat{S}_n = \frac{\sum_i X_i}{\sqrt{\sigma n}}$ has mean 0 and var 1.

1. Since the X_i are independent:

$$M_{\hat{S}_n}(s) = \Pi_{i=1}^n E[e^{\frac{X_i}{\sqrt{\sigma^2 n}}}] = (M_X(\frac{s}{\sigma\sqrt{n}}))^n$$

2. Find Taylor Expansion of $M_X(s)$ around $s = 0$.

$$M_X(s) = 1 + \frac{\sigma^2}{2}s^2 + o(s^2).$$

$$M_{\hat{S}_n}(s) = (M_X(\frac{s}{\sigma\sqrt{n}}))^n = (1 + \frac{s^2}{2n} + o(\frac{s^2}{\sigma^2 n}))^n.$$

3. As we take limits, only the second term remains!

$$\lim_{n\to\infty} \ln M_{\hat{S}_n}(s) = \lim_{n\to\infty} \ln(1 + \frac{s^2}{2n} + o(\frac{s^2}{\sigma^2 n}))^n$$

But this goes to $\to \frac{s^2}{2}, \quad \text{as } n \to \infty$. This is the log of the transform for a Standard Normal.

4. Thus for any x, $F_X(x) \to$CDF of a Standard Normal. I.e., $\hat{S}_n$ **converges in distribution** to $N(0, 1)$.

Strong Law of Large Numbers

Let $X_1, X_2, \ldots$ be a sequence of iid random variables with mean μ. Then

$$P(\lim_{n\to\infty} \frac{X_1 + \ldots + X_n}{n} = \mu) = 1.$$

Strong Law of Large Numbers

Let $X_1, X_2, ...$ be a sequence of iid random variables with mean μ. Then

$$P(\lim_{n\to\infty} \frac{X_1 + ... + X_n}{n} = \mu) = 1.$$

Think of the experiment being repeated n times. Then a basic outcome consists of $(x_1, x_2, ...x_n)$. The Strong Law says that as $n \to \infty$, all the probability is concentrated on those sequences that converge to μ.

Strong Law of Large Numbers

Let $X_1, X_2, \ldots$ be a sequence of iid random variables with mean μ. Then

$$P(\lim_{n \to \infty} \frac{X_1 + \ldots + X_n}{n} = \mu) = 1.$$

Think of the experiment being repeated n times. Then a basic outcome consists of $(x_1, x_2, \ldots x_n)$. The Strong Law says that as $n \to \infty$, all the probability is concentrated on those sequences that converge to μ.
We say that M_n **converges to** μ **almost surely**.

What's the difference between the Weak and Strong Laws?

If $X_1, ..., X_n$ are iid random variables with mean μ and $M_n = \frac{1}{n}\sum_i X_i$:

Weak Law of Large Numbers

For every $\epsilon > 0$:

$$\lim_{n\to\infty} P(|M_n - \mu| \geq \epsilon) = 0$$

Strong Law of Large Numbers

$$P(\lim_{n\to\infty} M_n - \mu = 0) = 1.$$

Point of difference: For some $\epsilon > 0$, how often is $|M_n - \mu| > \epsilon$?

What's the difference between the Weak and Strong Laws?

If $X_1, \ldots, X_n$ are iid random variables with mean μ and $M_n = \frac{1}{n}\sum_i X_i$:

Weak Law of Large Numbers

For every $\epsilon > 0$:

$$\lim_{n\to\infty} P(|M_n - \mu| \geq \epsilon) = 0$$

Strong Law of Large Numbers

$$P(\lim_{n\to\infty} M_n - \mu = 0) = 1.$$

Point of difference: For some $\epsilon > 0$, how often is $|M_n - \mu| > \epsilon$?

1. Weak Law:

$$\frac{\text{Number of times } |M_n - \mu| > \epsilon}{n} \to 0$$

But this still allows for the numerator to be unbounded with n. (For example $\ln n$.)

What's the difference between the Weak and Strong Laws?

If $X_1, \ldots, X_n$ are iid random variables with mean μ and $M_n = \frac{1}{n}\sum_i X_i$:

Weak Law of Large Numbers

For every $\epsilon > 0$:

$$\lim_{n\to\infty} P(|M_n - \mu| \geq \epsilon) = 0$$

Strong Law of Large Numbers

$$P(\lim_{n\to\infty} M_n - \mu = 0) = 1.$$

Point of difference: For some $\epsilon > 0$, how often is $|M_n - \mu| > \epsilon$?

1. Weak Law:

$$\frac{\text{Number of times } |M_n - \mu| > \epsilon}{n} \to 0$$

   But this still allows for the numerator to be unbounded with n. (For example $\ln n$.)

2. Strong Law: Has to be a finite number of times for $\lim_{n\to\infty} M_n - \mu$ to be zero.

Example

Suppose time is in discrete units, i.e., $1, 2, ...,$ and $Y_n = 1$ if there is an arrival at time n and $Y_n = 0$ otherwise.
Define $I_k = \{2^k, 2^k + 1, ..., 2^{k+1} - 1\}$ so that:

$$I_1 = \{1\},\ I_2 = \{2, 3\},\ I_3 = \{4, 5, 6, 7\} \quad \text{etc.}$$

Suppose that during each interval I_k we have exactly one arrival, and that arrival is equally likely to be at any time in that interval.

$P(Y_1 = 1) = 1,\ P(Y_2 = 1) = P(Y_3 = 1) = 0.5,\ P(Y_4 = 1) = ... = P(Y_7 = 1) = 0.2$

Example

Suppose time is in discrete units, i.e., $1, 2, ...,$ and $Y_n = 1$ if there is an arrival at time n and $Y_n = 0$ otherwise.
Define $I_k = \{2^k, 2^k + 1, ..., 2^{k+1} - 1\}$ so that:

$$I_1 = \{1\},\ I_2 = \{2, 3\},\ I_3 = \{4, 5, 6, 7\} \quad \text{etc.}$$

Suppose that during each interval I_k we have exactly one arrival, and that arrival is equally likely to be at any time in that interval.

$P(Y_1 = 1) = 1,\ P(Y_2 = 1) = P(Y_3 = 1) = 0.5,\ P(Y_4 = 1) = ... = P(Y_7 = 1) = 0.2$

Then $P(Y_n = 1) = \frac{1}{2^k}$ if $n \in I_k$.

$$\lim_{n\to\infty} P(Y_n = 1) = \lim_{k\to\infty} \frac{1}{2^k} = 0.$$

Thus, Y_n converges to 0 in probability.

Example

Suppose time is in discrete units, i.e., $1, 2, ...,$ and $Y_n = 1$ if there is an arrival at time n and $Y_n = 0$ otherwise.
Define $I_k = \{2^k, 2^k + 1, ..., 2^{k+1} - 1\}$ so that:

$$I_1 = \{1\},\ I_2 = \{2, 3\},\ I_3 = \{4, 5, 6, 7\} \quad \text{etc.}$$

Suppose that during each interval I_k we have exactly one arrival, and that arrival is equally likely to be at any time in that interval.

$P(Y_1 = 1) = 1,\ P(Y_2 = 1) = P(Y_3 = 1) = 0.5,\ P(Y_4 = 1) = ... = P(Y_7 = 1) = 0.2$

Then $P(Y_n = 1) = \frac{1}{2^k}$ if $n \in I_k$.

$$\lim_{n\to\infty} P(Y_n = 1) = \lim_{k\to\infty} \frac{1}{2^k} = 0.$$

Thus, Y_n converges to 0 in probability.
However, $P(\lim_{n\to\infty} Y_n = 0) \neq 0$, since given any finite n there are certain to be an infinite number of arrivals after n.
So Y_n does not converge almost surely.

What have learned so far?

1. Discrete Sample Spaces and the Axioms of Probability: Conditioning, Independence, Bayes Rule, Counting

What have learned so far?

1. Discrete Sample Spaces and the Axioms of Probability: Conditioning, Independence, Bayes Rule, Counting
2. Discrete Random Variables: Pmfs, Expectation and Variance, Conditioning, Independence, Bayes Rule.

What have learned so far?

1. Discrete Sample Spaces and the Axioms of Probability: Conditioning, Independence, Bayes Rule, Counting
2. Discrete Random Variables: Pmfs, Expectation and Variance, Conditioning, Independence, Bayes Rule.
3. Continuous Random Variables: Density Functions, CDFS, Conditioning, Independence, Bayes Rule.

What have learned so far?

1. Discrete Sample Spaces and the Axioms of Probability: Conditioning, Independence, Bayes Rule, Counting
2. Discrete Random Variables: Pmfs, Expectation and Variance, Conditioning, Independence, Bayes Rule.
3. Continuous Random Variables: Density Functions, CDFS, Conditioning, Independence, Bayes Rule.
4. Further Topics: Derived Distributions, Covariance, Correation, Iterated Expectations, Transforms.

What have learned so far?

1. Discrete Sample Spaces and the Axioms of Probability: Conditioning, Independence, Bayes Rule, Counting
2. Discrete Random Variables: Pmfs, Expectation and Variance, Conditioning, Independence, Bayes Rule.
3. Continuous Random Variables: Density Functions, CDFS, Conditioning, Independence, Bayes Rule.
4. Further Topics: Derived Distributions, Covariance, Correation, Iterated Expectations, Transforms.
5. Inequalities, WLLN, SLLN, CLT.

What have learned so far?

1. Discrete Sample Spaces and the Axioms of Probability: Conditioning, Independence, Bayes Rule, Counting
2. Discrete Random Variables: Pmfs, Expectation and Variance, Conditioning, Independence, Bayes Rule.
3. Continuous Random Variables: Density Functions, CDFS, Conditioning, Independence, Bayes Rule.
4. Further Topics: Derived Distributions, Covariance, Correation, Iterated Expectations, Transforms.
5. Inequalities, WLLN, SLLN, CLT.

What more is there??

What more is there??

Just applications of what we have learned so far...

1. Stochastic Processes: Basic Outcomes of the experiment are tied to time.
   - Arrival Processes: Discrete time (Bernoulli), Continuous time (Poisson)
   - Markov Chains: Discrete time, Continuous time

What more is there??

Just applications of what we have learned so far...

1. Stochastic Processes: Basic Outcomes of the experiment are tied to time.
   - Arrival Processes: Discrete time (Bernoulli), Continuous time (Poisson)
   - Markov Chains: Discrete time, Continuous time
2. Estimation: Bayesian v/s Classical

Stochastic Processes

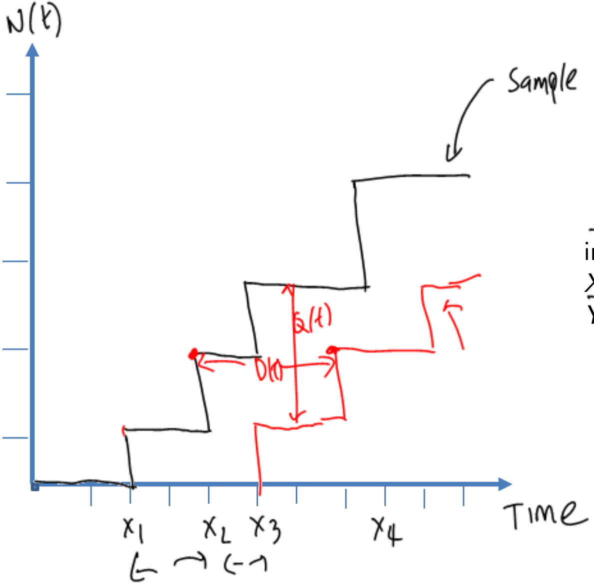
Suppose our Experiment is to measure $N(t)$ the number of people who leave a bookstore in the interval $[0, T]$. $N(t)$ is a random variable that depends on time. We refer to such an experiment as a "process".

Stochastic Processes

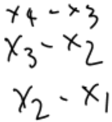
Suppose our Experiment is to measure $N(t)$ the number of people who leave a bookstore in the interval $[0, T]$. $N(t)$ is a random variable that depends on time. We refer to such an experiment as a "process". Examples:

1. The price fluctuations of a stock: How do the future prices depend on past prices?
2. The behavior of a queue over time: What happens to it if we add more cashiers?
3. The sequence of failures times of a machine: Is there a pattern to these failures? Correlation to the failures of other machines in the factory?
4. etc

Arrival Processes

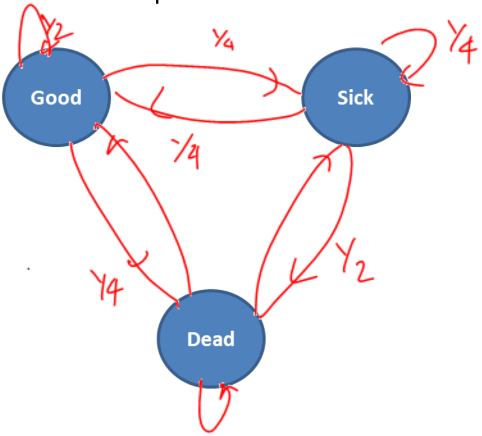


$N(t)$: number of arrivals in $[0, t]$
X_k: time of k^{th} arrival
Y_k: k^{th} interarrival time.



Markov Processes

- Future **depends** on past. However, future value depends only the current state of the system.
- Example: Machine Health



	Good	S	D
Good	1/2	1/4	1/4
Sick			
Dead			

Bernoulli Process

An arrival process comprised of a sequence of coin flips: $X_1, X_2, \ldots$ where the X_i are iid Bernoulli rv's with probability of success p.

Bernoulli Process

An arrival process comprised of a sequence of coin flips: $X_1, X_2, \dots$ where the X_i are iid Bernoulli rv's with probability of success p.

1. Let $N(t)$ be the number of arrivals in t time slots. Then, $N(t) =$

Bernoulli Process

An arrival process comprised of a sequence of coin flips: $X_1, X_2, \ldots$ where the X_i are iid Bernoulli rv's with probability of success p.

1. Let $N(t)$ be the number of arrivals in t time slots. Then, $N(t) = \sum_i X_i$ is Binomial.

Bernoulli Process

An arrival process comprised of a sequence of coin flips: $X_1, X_2, \ldots$ where the X_i are iid Bernoulli rv's with probability of success p.

1. Let $N(t)$ be the number of arrivals in t time slots. Then, $N(t) = \sum_i X_i$ is Binomial.
2. Let Y_k be the time of the k^th arrival.
   Then Y_1 is the time to the first head i.e. $Y_1 \sim$

Bernoulli Process

An arrival process comprised of a sequence of coin flips: $X_1, X_2, \ldots$ where the X_i are iid Bernoulli rv's with probability of success p.

1. Let $N(t)$ be the number of arrivals in t time slots. Then, $N(t) = \sum_i X_i$ is Binomial.
2. Let Y_k be the time of the $k^t h$ arrival.
   Then Y_1 is the time to the first head i.e. $Y_1 \sim$ Geometric with parameter p.

Bernoulli Process

An arrival process comprised of a sequence of coin flips: $X_1, X_2, \ldots$ where the X_i are iid Bernoulli rv's with probability of success p.

1. Let $N(t)$ be the number of arrivals in t time slots. Then, $N(t) = \sum_i X_i$ is Binomial.
2. Let Y_k be the time of the $k^t h$ arrival.
   Then Y_1 is the time to the first head i.e. $Y_1 \sim$ Geometric with parameter p.
3. $T_1 = Y_1$ is the time to the first arrival. $T_2 = Y_2 - Y_1$ is the time between the first and second arrivals. Since coin flipping is memoryless, $Y_2 - Y_1$ is also geometric and independent of Y_1.

Bernoulli Process

An arrival process comprised of a sequence of coin flips: $X_1, X_2, ...$ where the X_i are iid Bernoulli rv's with probability of success p.

1. Let $N(t)$ be the number of arrivals in t time slots. Then, $N(t) = \sum_i X_i$ is Binomial.
2. Let Y_k be the time of the k^th arrival.
   Then Y_1 is the time to the first head i.e. $Y_1 \sim$ Geometric with parameter p.
3. $T_1 = Y_1$ is the time to the first arrival. $T_2 = Y_2 - Y_1$ is the time between the first and second arrivals. Since coin flipping is memoryless, $Y_2 - Y_1$ is also geometric and independent of Y_1.

$T_1, T_2, T_3, T_4 \ldots$

Definition 2

A sequence of iid geometric random variables with parameter p.

Example: Weather

The number of days between rainy days has been observed to be geometric with parameter p and independent of any other period. What is the probability that it will rain on day 8 and 12 of the next month?

Example: Weather

The number of days between rainy days has been observed to be geometric with parameter p and independent of any other period. What is the probability that it will rain on day 8 and 12 of the next month?

View rainy days as "arrivals". Then they must form a Bernoulli Process since the inter-arrival times are iid Geometric.

Example: Weather

The number of days between rainy days has been observed to be geometric with parameter p and independent of any other period. What is the probability that it will rain on day 8 and 12 of the next month?
View rainy days as "arrivals". Then they must form a Bernoulli Process since the inter-arrival times are iid Geometric.
So both probability that it will rain on any day is p. Answer p^2.

Distribution of the Time of the k^{th} Arrival

Y_k is the sum of the first k interarrival times.
The interarrival times are geometric rvs.
(Geometric Mean: $1/p$; Var: $(1-p)/p^2$.)

$$\boxed{E[Y_k] = \frac{k}{p}}$$

$$\boxed{var(Y_k) = \frac{k(1-p)}{p^2}}$$

Distribution of the Time of the k^{th} Arrival

Y_k is the sum of the first k interarrival times.
The interarrival times are geometric rvs.
(Geometric Mean: $1/p$; Var: $(1-p)/p^2$.)

$$E[Y_k] = \frac{k}{p}$$

$$var(Y_k) = \frac{k(1-p)}{p^2}$$

In the first $t-1$ time units we have to had exactly $k-1$ arrivals.
The t^{th} time unit must register an arrival.

In t-1 slots exactly k-1 arrivals

$$P(Y_k = t) = p\binom{t-1}{k-1}p^{k-1}(1-p)^{t-k}$$

Distribution of the Time of the k^{th} Arrival

Y_k is the sum of the first k interarrival times.
The interarrival times are geometric rvs.
(Geometric Mean: $1/p$; Var: $(1-p)/p^2$.)

$$\boxed{E[Y_k] = \qquad\qquad}$$

$$\boxed{var(Y_k) = \qquad\qquad}$$

In the first $t-1$ time units we have to had exactly $k-1$ arrivals.
The t^{th} time unit must register an arrival.

$$P(Y_k = t) =$$

Distribution of the Time of the k^{th} Arrival

Y_k is the sum of the first k interarrival times.
The interarrival times are geometric rvs.
(Geometric Mean: $1/p$; Var: $(1-p)/p^2$.)

$$\boxed{E[Y_k] = \qquad}$$

$$\boxed{var(Y_k) = \qquad}$$

In the first $t-1$ time units we have to had exactly $k-1$ arrivals.
The t^{th} time unit must register an arrival.

$$\boxed{P(Y_k = t) = \binom{t-1}{k-1} p^{k-1}(1-p)^{t-k} p = \binom{t-1}{k-1} p^k (1-p)^{t-k}}$$

Called Pascal Distribution of order k.

Example: Priority System

A packet arrives at a network device at the beginning of each time slot with probability p. It takes 1 slot to process. A Busy/Idle Period is a maximal string of busy/idle slots. Find the distributions of:

BBI BII BI BI BBBI

- Time to first busy slot:

Example: Priority System

A packet arrives at a network device at the beginning of each time slot with probability p. It takes 1 slot to process. A Busy/Idle Period is a maximal string of busy/idle slots. Find the distributions of:

- Time to first busy slot:
  Geometric with p
- Time to first idle slot:

Example: Priority System

A packet arrives at a network device at the beginning of each time slot with probability p. It takes 1 slot to process. A Busy/Idle Period is a maximal string of busy/idle slots. Find the distributions of:

- Time to first busy slot:
  Geometric with p
- Time to first idle slot:
  Geometric with $1 - p$
- Length of first busy period:

Example: Priority System

A packet arrives at a network device at the beginning of each time slot with probability p. It takes 1 slot to process. A Busy/Idle Period is a maximal string of busy/idle slots. Find the distributions of:

- Time to first busy slot:
  Geometric with p
- Time to first idle slot:
  Geometric with $1 - p$
- Length of first busy period:
  Geometric with $1 - p$
- Length of the first idle period:

Example: Priority System

A packet arrives at a network device at the beginning of each time slot with probability p. It takes 1 slot to process. A Busy/Idle Period is a maximal string of busy/idle slots. Find the distributions of:

- Time to first busy slot: Geometric with p
- Time to first idle slot: Geometric with $1 - p$
- Length of first busy period: Geometric with $1 - p$
- Length of the first idle period: Geometric with with p

Example: Two A's

Bob gets an A in his quizzes independently with probability p.

1. P(He gets exactly two A's in the next six quizzes)=

Example: Two A's

Bob gets an A in his quizzes independently with probability p.

1. P(He gets exactly two A's in the next six quizzes)= Binomial $k = 2$, $n = 6$.
2. E[quizzes he must take before he has gotten two As]=

Example: Two A's

Bob gets an A in his quizzes independently with probability p.

1. P(He gets exactly two A's in the next six quizzes)$=$ Binomial $k = 2$, $n = 6$.
2. E[quizzes he must take before he has gotten two As]$=$ E[2nd arrival time] $=$

Example: Two A's

Bob gets an A in his quizzes independently with probability p.

1. P(He gets exactly two A's in the next six quizzes)= Binomial $k = 2,\ n = 6$.
2. E[quizzes he must take before he has gotten two As]= E[2nd arrival time] $= 2(1/p)$
3. P[1st and 2nd time he gets an A is on the 8th and 9th Quiz]=

Example: Two A's

Bob gets an A in his quizzes independently with probability p.

1. P(He gets exactly two A's in the next six quizzes)= Binomial $k = 2,\ n = 6$.
2. E[quizzes he must take before he has gotten two As]= E[2nd arrival time] $= 2(1/p)$
3. P[1st and 2nd time he gets an A is on the 8th and 9th Quiz]=

$$P(Y_1 = 8, Y_2 = 9) = P(Y_1 = 8)P(Y_2 - Y_1 = 1) = (1-p)^7 p^2$$

Arriving at a Random Time

Alice has been practicing probability problems. She gets one right with probability p independent of how she has done on the others. Bob walks in and watches Alice do a problem correctly. What is the pmf of the Alice's "success period" which includes this problem?

WRLW(L R)WRRRW

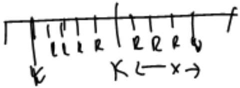
Arriving at a Random Time

Alice has been practicing probability problems. She gets one right with probability p independent of how she has done on the others. Bob walks in and watches Alice do a problem correctly. What is the pmf of the Alice's "success period" which includes this problem?

- $K-1$: # problems Alice got correct before Bob walked into the room.

Arriving at a Random Time

Alice has been practicing probability problems. She gets one right with probability p independent of how she has done on the others. Bob walks in and watches Alice do a problem correctly. What is the pmf of the Alice's "success period" which includes this problem?

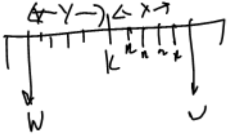


- $K-1$: # problems Alice got correct before Bob walked into the room.
- Over all time the success period is geometric with $1-p$.
- X: num problems to first failure **after** K.
- Y: num problems to first failure **before** K

Arriving at a Random Time

Alice has been practicing probability problems. She gets one right with probability p independent of how she has done on the others. Bob walks in and watches Alice do a problem correctly. What is the pmf of the Alice's "success period" which includes this problem?

- $K-1$: # problems Alice got correct before Bob walked into the room.
- Over all time the success period is geometric with $1-p$.
- X: num problems to first failure **after** K.
- Y: num problems to first failure **before** K
- X, Y independent and geometric with parameter $1-p$.



Arriving at a Random Time

Alice has been practicing probability problems. She gets one right with probability p independent of how she has done on the others. Bob walks in and watches Alice do a problem correctly. What is the pmf of the Alice's "success period" which includes this problem?

- $K-1$: # problems Alice got correct before Bob walked into the room.
- Over all time the success period is geometric with $1-p$.
- X: num problems to first failure **after** K.
- Y: num problems to first failure **before** K
- X, Y independent and geometric with parameter $1-p$.
- We want pmf of $X+Y-1$.
- $X+Y$ is Pascal order 2.

Random Incidence

1. The red and blue bus lines take you where you want to go. The red bus arrives on the hour every hour. The blue bus arrives at one minute past the hour, every hour. You show up at some random time. What is the probability you will take the red bus?

Random Incidence

1. The red and blue bus lines take you where you want to go. The red bus arrives on the hour every hour. The blue bus arrives at one minute past the hour, every hour. You show up at some random time. What is the probability you will take the red bus?
2. Interarrival times: After a bus has arrived, the next bus arrives in 10mins with prob .9 and in 100 min with prob .1. You arrive at a random time. What is your expected wait? The overall expected wait time is 19 mins.

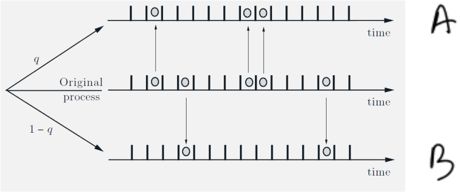
Random Incidence

1. The red and blue bus lines take you where you want to go. The red bus arrives on the hour every hour. The blue bus arrives at one minute past the hour, every hour. You show up at some random time. What is the probability you will take the red bus?
2. Interarrival times: After a bus has arrived, the next bus arrives in 10mins with prob .9 and in 100 min with prob .1. You arrive at a random time. What is your expected wait? The overall expected wait time is 19 mins.
   Suppose N busses have arrived. Then for large N, the total time $T \approx (9/10)5N + (1/10)100N = 14.5N$. Prob that you arrive in a 100 min interval is $10/14.5$ and the prob you arrive in a short interval is $4.5/14.5$.

Random Incidence

1. The red and blue bus lines take you where you want to go. The red bus arrives on the hour every hour. The blue bus arrives at one minute past the hour, every hour. You show up at some random time. What is the probability you will take the red bus?
2. Interarrival times: After a bus has arrived, the next bus arrives in 10mins with prob .9 and in 100 min with prob .1. You arrive at a random time. What is your expected wait? The overall expected wait time is 19 mins.
   Suppose N busses have arrived. Then for large N, the total time $T \approx (9/10)5N + (1/10)100N = 14.5N$. Prob that you arrive in a 100 min interval is $10/14.5$ and the prob you arrive in a short interval is $4.5/14.5$. Assuming that you are equally likely to be in anywhere in the interval you arrive in...
   Expected wait $= 50(0.69) + 5(0.31) \approx 35$ mins.

Splitting Bernoulli Processes

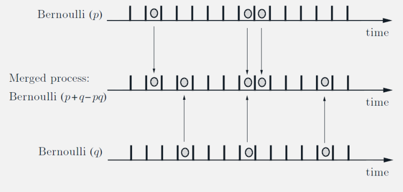


For each arrival, route it to A with prob q and to B with prob $1 - q$. Routing decisions are independent.
A_i: arrival to A at time slot i, and B_i: arrival to B at slot i.

$$P(A_i) = \begin{cases} 1, & \text{if } X_i = 1 \text{ and route to } A; \\ 0, & \text{otherwise.} \end{cases}$$

So, A_i is Bernoulli with prob pq and B_i is Bernoulli with probability $p(1 - q)$.

Merging Bernoulli Processes



A and B are independent Bernoulli Processes, with prob p_A and p_B respectively. Construct the arrival process C as follows:
$C_i = 1$ if $A_i = 1$ or $B_i = 1$ (includes if both have arrivals) and
$C_i = 0$ otherwise.
Then $P(C_i = 0) = (1 - p_A)(1 - p_B)$ and
$P(C_i = 1) = p_C = 1 - (1 - p_A - p_B + p_A p_B) = p_A + p_B - p_A p_B$.
Thus C is Bernoulli Process with probability p_C.

Example: Sum of a geometric number of geometric rvs

$$Y = \sum_{i=1}^{N} X_i$$

X_i's geometric with p and N is geometric with q
$X_1, X_2, ...$ interarrival times of a Bernoulli process(p).

Example: Sum of a geometric number of geometric rvs

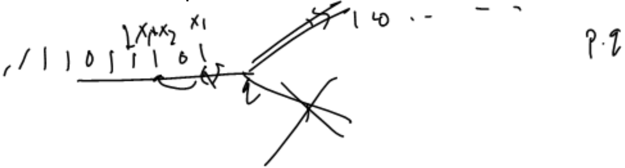
$$Y = \sum_{i=1}^{N} X_i$$

X_i's geometric with p and N is geometric with q

$X_1, X_2, \ldots$ interarrival times of a Bernoulli process(p).

$X_1, X_1 + X_2, X_1 + X_2 + X_3,$: arrival times of a Bernoulli process(p)

Construct a split Bernoulli Process as follows:



Example: Sum of a geometric number of geometric rvs

$$Y = \sum_{i=1}^{N} X_i$$

X_i's geometric with p and N is geometric with q
$X_1, X_2, ...$ interarrival times of a Bernoulli process(p).
$X_1, X_1 + X_2, X_1 + X_2 + X_3,$: arrival times of a Bernoulli process(p)
Construct a split Bernoulli Process as follows:

Each arrival is rejected with prob $1 - q$.
The split Bernoulli has parameter pq. The number of arrivals to the first acceptance is geometric with parameter q. Therefore

$$Y = X_1 + ... + X_N$$

is geometric with parameter pq.

Enjoy Spring Break